Scientifico! like Dad
On the Intergenerational Transmission of STEM Education in Italy*

Diana Chise† Margherita Fort§ Chiara Monfardini§

DRAFT PREPARED FOR THE XXXIII AIEL CONFERENCE
PRELIMINARY AND INCOMPLETE

September 10, 2018

Abstract
We provide novel evidence on the existence and the extent of intergenerational transmission of STEM (Science, Technology, Engineering and Mathematics) education using a recent large administrative dataset of Italian graduates obtained from the AlmaLaurea database. We look at the parental influence on two STEM educational outcomes (High School and University degree completion) and exploit the repeated choice over time to account for student’s unobserved heterogeneity. We allow for gendered effects according to both the parent and the student’s gender and for full interaction among father and mother education. We find that having a parent with a STEM degree has a strong positive effect on the probability that the student pursues a STEM field degree at both educational stages. Fathers are more influential than mothers on the University outcome, but not on high school outcome. Females and males are similarly affected by the education of each parent at University, while females are more affected than males at high school. The fixed effect strategy results show that mothers matter more for high school than for University choices, while the opposite is true for fathers and that these effects are driven by same gender interaction (father-son, mother-daughter).

JEL-Codes: J16, J24, I24

Keywords: gender, intergenerational transmission, parents, STEM

*This paper borrows from the LMEC master thesis at the University of Bologna of D. Chise under the supervision of C. Monfardini that circulated under the title “Intergenerational transmission of STEM educational choices in Italy: a gender perspective”. “Scientifico!” it is also the title of a performance by A. Libertini and V. Nah (company: Piccoli Principi-Italy/ Athénor -France/ Teatro all’improvviso-Italy), who bring on stage a special science lesson for young children. We are very grateful to D. Cristofori and S. Ghiselli for their help in clarifying the many institutional and administrative details of the AlmaLaurea data set. We acknowledge the outstanding work of F. Meluzzi as research assistant. Finally, we are grateful to M. Anelli, M. De Paola and G. Moroni and other seminar participants at the Ninth International Workshop on Applied Economics of Education (IX IWAEE), at the University of Bologna for very valuable comments and suggestions. Authors’ email addresses: diana.chise@unibo.it; margherita.fort@unibo.it; chiara.monfardini@unibo.it.

§University of Bologna, Department of Economics, and IZA. Fort is also affiliated with CESifo. Monfardini is also affiliated with CHILD-CCA.

†European Central Bank, Market Operations Directorate; prospective PhD candidate at Trinity College Dublin.
1 Introduction

The demand for professional and associate professional occupations in the fields of Science, Technology, Engineering and Mathematics (STEM) is expected to grow by 13% and 7% respectively in the period 2015-2025 in Europe, as opposed to a predicted 3% increase in employment for all occupations (EU Skills Panorama 2014, 2015). In the US, the projected growth of STEM occupations is even more rapid, with the STEM mathematical science occupations group being expected to grow by 28% over the 2014-2024 period, compared with the 6.5% predicted growth for all occupations (Fayer et al., 2015). However, the future supply of STEM related skills is feared to result insufficient worldwide. This shortage of supply is partly due to the under-representation of women in these fields, which is a reason of growing concern for policymakers and social scientists (European Commission, 2015; UNESCO, 2017; Kahan and Ginther, 2017). Although women have globally closed the gap in terms of primary and secondary education level, and currently represent the majority of students enrolled in higher education, horizontal segregation in the field of studies in secondary and higher education still persists (UNESCO, 2017) and STEM occupations and jobs are heavily male dominated (National Science Foundation, 2017; Kahan and Ginther, 2017). Gender differences in sorting into STEM education and jobs are likely to have implications on the average wage gender gap. Indeed, STEM high school education has been documented by Black et al. (2015) to be an important predictor of labour market success, as measured by wages and occupational upgrading in mid-life, and STEM occupations are recognized to be more rewarded in the job market than NON-STEM ones (Goos et al., 2013; Fayer et al., 2015).¹

Recent studies have empirically explored the role of STEM education as a driver of economic performance at the aggregate level. According to Ray (2015), the presence of STEM graduates stimulated economic growth and innovation in the USA.² There is also evidence that STEM graduates generates positive externalities on the society at large, improving the

¹According to this report the national average wage for all STEM occupations in 2015 was nearly double the national average wage for non-STEM occupations in the US.
²She finds that the share of STEM graduates has a positive impact on the level and growth of real GDP per capita and on the number of patents in the 50 US states and the District of Columbia between 1990 and 2011.
The ability to attract future generations of STEM students, to stimulate the supply of STEM professionals and to reduce its gender unbalance rests on the understanding of the forces that drive self-selection into STEM fields. This has very recently motivated a growing literature exploring the determinants of STEM enrollment and of its gender gap. The possible factors standing as influential for the field choice range from peer effects (Anelli and Peri, forthcoming) to sibling gender composition (Brenoe, 2017; Oguzoglu and Ozbeklik, 2016), previous science exposure (De Philippis, 2017; Gottfried and Bozick, 2016), information bareers (Barone et al., 2017), beliefs and expectations (Wiswall and Zafar, 2015), and socio-cultural background measured by conservative political and religious attitudes (Grossmann et al., 2016). In parallel, researchers have been analysing the drivers of STEM gender gap in education. Card and Abigail (2017) and Granato (2018) find that the gender gap in STEM university entry and graduation is substantially explained by high school STEM readiness and scientific content. On the contrary, the proportion of high school female peers is estimated to widen gender differences in STEM higher education (Brenoe and Zolitz, 2018).

Despite family background is widely recognized as as a major determinant of educational attainments in developed countries, its role in shaping STEM education has been under-explored in the literature and the evidence on the relation between parental characteristics and STEM fields choice or graduation is very scarce. To the best of our knowledge, the only exception is Granato (2018) which investigates the early determinants of the STEM gender gap in Italy. She finds that parental social status and education are positively associated with the probability of achieving a STEM degree and of choosing a high school with greater maths content.

On the other side, the vast literature on the intergenerational transmission of inequality

---

3Based on data from the American Community Survey 2009-2011, this paper shows that College graduates in both STEM and non-STEM fields create positive externalities on wages of workers of the same metropolitan area, but STEM graduates create much larger externalities.

4A related literature focuses on the gender gap in math performance (Contini et al., 2017), emphasizing cultural explanations (Guiso et al., 2008; Nollenberger et al., 2016), and the impact of teachers’ gender stereotypes (Carlan, 2018). Other papers reveal the existence of gender differences in persistence in STEM education and analyse their sources (Ehrenberg, 2010; Fischer, 2017; Griffith, 2010).

5Some related results are contained in Artmann et al. (2018), who study the effect of the field of study on future family outcomes in the Netherlands, but produce also estimates of the effect of parental degree on students’ academic enrollment in Medicine and International Business Economics.
in education has thoroughly investigated the vertical transmission of years of schooling (see Björklund and Salvanes (2010); Black and Devereux (2011); Holmlund et al. (2011) for exhaustive reviews and Amin et al. (2015) for the importance of allowing for heterogeneous effects according to parents’ and children gender). To the best of our knowledge, no study so far has addressed explicitly the issue of linking the STEM educational attainments of parents and children, despite the educational expansion over past years increased the importance of horizontal differentiation between fields of study (Triventi, 2013).

In this paper we aim at filling this gap. Exploiting a recent, high quality dataset on a large portion of the population of Italian graduates (2016 cohort) we provide novel and up-to-date evidence on the existence and the extent of intergenerational transmission of STEM educational outcomes, focusing on gendered effects according to both parents’ and students’ gender. The case study of Italy we analyse is particularly interesting since the Italian society is characterized by high levels of both intergenerational persistence of educational attainment (Checchi et al., 2013) and gender inequality in most socio-economic domains (World Economic Forum, 2017). Since the beginning of the nineties, when Italian female graduates have reached their males counterpart, female enrollment has witnessed a continuous increase and stabilized around 60% of the total number of graduates in the recent years. Despite representing more than half of the graduate population, however, women tend to earn lower wages with respect to men. Piazzalunga (2018) reports an overall gender gap estimated by EUROSTAT (2017) of about 44%. One of the explanation of this disadvantage can be found in females’ under-representation in STEM education, as argued by Anelli and Peri (2015). In our sample approximately 23% of students graduate in STEM fields but this share decreases to 16% among females while it is about 35% among males. A distinguished feature of our study is the use of information on students’ field of qualification at two points in time: at completion of high school and of university, respectively. Exploiting the non-negligible variability in field of studies across the two educational stages, we resort to a fixed effect strategy and are able to identify changes across time in intergenerational coefficients controlling for any unobserved student level time-constant characteristics. This is a non trivial approach in intergenerational studies which allow us to explore whether the influence of parents increases or diminishes at the university stage with respect to the high school one.
A key ingredient of our research is represented by the data we build on parents’ education. The existing Almalaurea data contain classified information only on the level of parents’ education. We had access to uncoded, text-type, information on the parents degree type contained in the Almalaurea survey and identified the field of study of the parents with a tedious process of extraction and elaboration.\(^6\)

Our findings document a sizeable role of parents’ STEM field degree and heterogenous effects according to the gender of the parent, the gender of the student and the educational stage (high school or university).

The estimated intergenerational persistence in field of study at University it is not totally attributable to the student field choice in high school and to parental influence on this intermediate choice. Fathers exert a larger influence than mothers do on University completion, not on high school completion, for which both parents appear equally important. Females and males students are similarly affected by each parent at University, while females are more sensitive to the influence of each parent at high school. When we allow for interdependent parent’s education effects we spot non-trivial interactions confirming the prominent role of father’s field of education with respect to mother’s and revealing a mother-daughter relationship. The fixed effects estimates enlight a differential role of STEM graduate parents at University with respect to high school: the importance of father increases across different educational stages, while that of mother diminishes. These effects are found to be driven by same gender interaction (father-son, mother-daughter).

Previous waves of the Almalaurea dataset have been recently independently used by Granato (2018) to explore the role of early influences on the STEM gender gap in Italy. While abstaining from an intergenerational trasmission perspective, her paper embedds cross sectional estimates of a positive association between parental and children STEM education, at both university and high school stages, which are consistent with our findings. These OLS intergenerational parameters are also found to be stronger for fathers than for mothers, but -given the different focus of the paper - they are not differentiated by students’ gender or education of the other parent.

Our paper is the first to identify a large set of heterogeneous intergenerational parameters

\(^6\)Details on the coding of the data on parent’s education can be found in section 3.2.
and represents an important preliminary step to isolate causal links, highlighting that there is room to promote gender equality also within the household. The results that exploit the longitudinal dimension of the data are fully consistent with the results that do not control for unobserved student characteristics. Moreover, we provide evidence that our intergenerational correlations are not explained by the transmission of liberal profession (a mechanism which has been spotted by Aina and Nicoletti (2018) in Italy), and are robust to a number of sensitivity checks. A less conservative view of our results could thus interpret our estimates as causal. We discuss the issue in more detail in the paper.

The paper is structured as follows. Section 2 illustrates the data, the process of classification of the parents’ degree and describes the sample we use. Section 3 presents gendered results of the intergenerational analysis in STEM educational outcomes. Section 4 concludes.

2 The Almalaurea Dataset

2.1 Sample selection and description

This study relies on data from the AlmaLaurea\(^7\) survey on the 2016 cohort of graduates from Italian Universities (XIX Survey “Profilo dei Laureati”). 71 Universities participated in the survey, covering about 90% of all Italian graduates.\(^8\) The response rate was above 92%, with a total of 272,225 participating students.

We focus on Italian graduates. While foreign students who have completed their high school abroad are excluded from our sample, we include foreign students who completed

\(^7\)AlmaLaurea is an Italian Interuniversity Consortium established in 1994 with the objective of conducting statistical studies on the Italian University system. AlmaLaurea runs surveys annually on the Profile of the graduates (“Profilo dei Laureati”) and their Employment status after 1, 3 and 5 years (“Condizione occupazionale dei Laureati”).

high school in Italy as they are most likely to be permanent residents, hence comparable to Italian students. The Italian University system has been sizebly reformed after the Bologna Process (1999) (Ministerial Decree no. 509/99). The Almalaurea database includes students who graduated in two different system: degrees obtained prior to 1999 correspond to the old system and degrees obtained after 1999 correspond to the new one. To allow for a consistent classification of the degrees, we focus on students who enrolled and graduated under the most recent system (“nuovo ordinamento”) and drop the students who hold degrees obtained under the old system (i.e. students who hold a “Laurea vecchio ordinamento”). We retain both students who graduated from a 3-years cycle degree (i.e. “Laurea”) as well as students who graduated from a 5-years cycle degree (i.e. “Laurea magistrale a ciclo unico”) with no missing values on all covariates used in the empirical analysis (dropped missing values account for about 18% of the initial sample size). We end up with a large final sample of 155,603 records, out of which approximately 60% are female students. We use both administrative and survey data. Administrative data cover information on the students’ University degree, including field of study. Survey data cover information on high school degree (including field of study) and parental background. The parents’ University degree is self-reported and we coded it from text.

Table 1 reports descriptive statistics on our working sample. Among graduates, approximately 23% choose a STEM degree at University and 51% choose a STEM degree at high school. The majority of both fathers and mothers have completed at most secondary education (around 46% of the fathers and 51% of the mothers). Overall, the share of fathers with tertiary education is slightly larger than the share of mother with tertiary education (21% of fathers compared to 19% of mothers). Among parents who graduated from University, more mothers than fathers hold non-STEM degrees (16% of mothers versus 14% of fathers), while the opposite is true for STEM degrees (6% of fathers versus 4% of mothers). Students raised in families where both the father and the mother hold STEM degrees make up 1% of the graduates, whereas approximately 7% of them is raised in families where both parents hold a non-STEM degree. The largest share of students in our sample come from families where both parents have at most high school qualifications (approximately 29%), followed by families where both parents have junior high school qualifications or less (approximately
Table 2 looks at the raw association between parental field of study and the choice of a STEM field degree at University: we document a positive association as the probability that the child enrolls in a STEM field degree is higher when the parent also graduated from a STEM field compared to the corresponding probabilities conditional on other levels of parental education. Tables 3 and 4 look at this association by gender of the child. The probability that a male child enrolls in a STEM degree increases by approximately 17 p.p. when his father is also a STEM graduate compared to the situation when the father has a junior high school qualification (from 35% to 49%), but only by 8 p.p. when it is the mother to be a STEM graduate (from 35% to 44%). The probability that a female child chooses a STEM degree is again higher when the father or the mother are STEM graduates rather than holding a junior high school qualification, going from 13% to 30% in the case of the father and from 13% to 27% in the case of the mother. A much stronger association is detected between son’s and father’s field of study than between son’s and mother’s field of study, while no such pattern is noticed for daugthers.

Table 5 presents descriptive evidence on the persistence of STEM field choices among students: about 82% of those who choose a non-STEM degree at high-school, do so also at University (and the share reaches about 91% for female students). However, the persistence of field choice is substantially more limited when one considers students who choose a STEM degree at high school: only 35% of them chooses a STEM degree also at University and this fraction decreases to 26% for female students. As a consequence a substantial fraction of student graduating from a STEM degree at University does not have a STEM high school degree (21% in the pooled sample; 33% among females). Overall, about 43% of students make the same non-STEM field choice at high school and at University while only 18% get a STEM-field degree at both high school and at University. The transition from a STEM degree at high school to a non-STEM degree at University is more likely than the reverse type of transition (30% versus 5%). Notably the persistence of STEM-field choice is much higher among offsprings of STEM field graduates: among children of fathers with STEM qualification, 30% choose a STEM field both at high school and at the University. We exploit this non-negligible variability in field choice over time in our empirical analysis to control for
any unobserved student level factor fixed over time while examining how intergenerational persistence of field choice evolves between high-school and university.

In our sample there is some indication of assortative mating by education level but to a much limited degree by field of study. Table 6 shows that parents tend to have the same qualification and men are more likely to “marry down”, i.e. marry women with lower qualifications, but the picture is more blurred when one considers fields of education: about 7% of the students in the sample have both parents graduated in a non-STEM degree compared to about 2% with both parents graduating from a STEM degree.\(^9\)

### 2.2 Classification of STEM degrees

STEM skills supply is defined as degree awarded in Science, Technology, Engineering and Math at the tertiary level. However, no common and detailed definition of which fields of study consists STEM core disciplines is available. In this paper, we use of the definition provided by the EU Commission in 2015, based on Eurostat’s Classification of Fields of Education and Training (1999). This classification from Eurostat is in turn based on the International Standard Classification of Education (ISCED 1997) developed by Unesco. ISCED has been revised several times, with the latest revision in 2013 focusing exclusively on fields of education and training (ISCED-F), and was implemented for EU data collections starting from 2016. In order to adhere to the up-to-date classification in this paper, we updated the grouping of the STEM macro categories provided by the EU 2015 report accordingly. The equivalent STEM fields of study can be classified in the following 3 categories: (i) Natural sciences, mathematics and statistics; (ii) Information and Communication Technologies; (iii) Engineering, manufacturing and construction.\(^10\) Differently from EU, we have chosen to include also Architecture and building since in Italy it is usually part of the department of Engineering and has many similarities with civil engineering studies.\(^11\) Instead, consistently

---

\(^9\)Descriptive statistics are computed on the sample of students and we cannot identify siblings who share the same parents in the data. Recall that all the students included in the sample belong to the same cohort and that it is not very common -with the exception of twins- to have siblings born on the same year.

\(^10\)The original fields as indicated by the European Commission can be found in a report available on line at https://publications.europa.eu/en/publication-detail/-/publication/60500ed6-cbd5-11e5-a4b5-01aa75ed71a1/language-en

\(^11\)Most of the alternative STEM definitions include Architecture and so does EU Skills Panorama 2014 (2015) in its analytical highlight on STEM skills.
with EU’s decision, we exclude Health studies, even though some studies include them. This decision was taken based on considerations on the historical evolution of the perception of different jobs as well as the presence of women. While Science has historically been viewed as non communicative, research focused and highly technical occupation, Medicine has been associated with personal care involving direct human contact and with overall equal distribution of men and women.\textsuperscript{12} A detailed list of the disciplines that make up STEM fields can be found in Appendix A.1.

The Italian university system is organised in three cycles, according to the Bologna framework adopted in 1999: the main academic degrees are the Laurea (1st cycle) which correspond to Bachelor degree, the Laurea Magistrale (2nd cycle) which is equivalent to Master degree, and the Dottorato di Ricerca (3rd cycle) - PhD. In our study we take into account 1st and 2nd cycle degrees. These are further structured into “classes”, according to the general educational objectives they share and core learning activities they must include. The Italian Ministry of Education has updated the correspondence between the Italian “classes” and the ISCED-F 2013 classification in 2016. This enabled us to give a very precise and up-to-date definition of which courses can be labeled as STEM according to the EC guideline. A complete list of the Italian classes classified as STEM can be found in Appendix A.2.

As mentioned in the introduction, we consider -beside the university field of study- the field of the high school degree previously achieved by the student. This required a further classification into STEM/NON-STEM of the secondary education diplomas obtained by the students and opening access to university. We categorized as STEM scientific high schools and selected technical high schools based on high mathematical/technical content of the curriculum of studies.\textsuperscript{13} All the remaining high schools were categorized as NON-STEM.\textsuperscript{14}

Finally, crucial to our intergenerational analysis is information on the field of the university degree achieved by the parents of the 2016 cohort of graduates under scrutiny. It is here that things became more intricate. The Almalaurea dataset does not provide structured data

\textsuperscript{12} We thank E. Luppi for the useful comments and suggestions.
\textsuperscript{13} STEM high schools include: \textit{liceo scientifico, istituto tecnico industriale, tecnico per geometri, tecnico nautico, tecnico aeronautico.}
\textsuperscript{14} NON-STEM high schools include: \textit{liceo classico, liceo psico-socio-pedagogico or istituto magistrale, liceo linguistico, liceo artistico, istituto darte, istituto professionale, istituto tecnico-commerciale, istituto tecnico-agrario, istituto tecnico per periti aziendali, istituto tecnico femminile per i servizi sociali.}
on parent’s educational fields. In their questionnaire, students answer this question by typing directly their parent’s degree title in an empty field which does not contain restrictions on the content and does not undergo any cleaning or standardization by the Almalaurea team. Due to the resulting large heterogeneity in inserted degree’ titles, a standardized classification was unavailable, and we classified the degrees into STEM/NON-STEM through an own-built procedure we describe in Appendix B.

3 Empirical analysis

We study intergenerational persistence in field of qualification at University (Block 1) and at high school (Block 2). More specifically, we start by estimating with OLS the parameters of equation (1) and equation (2) below.\textsuperscript{15}

\[ STEM_{i2} = \alpha_0 + \sum_f \alpha_f F Edu_f + \sum_m \alpha_m M Edu_m + \alpha' X_i + \varepsilon_{i2} \]  

\[ STEM_{i1} = \lambda_0 + \sum_f \lambda_f F Edu_f + \sum_m \lambda_m M Edu_m + \lambda' X_i + \varepsilon_{i1} \]  

where

(i) \( STEM_{i2} \) denotes a dummy taking the value 1 if student \( i \) graduates from a STEM field at University and 0 otherwise and \( STEM_{i1} \) denotes the corresponding variable that describes the field of graduation for the same student at high school;

(ii) \( F Edu_f, M Edu_m \ f, m \in \{ 1, 2, 3, 4 \} \) are dummy variables denoting the qualification level of fathers and mothers respectively, where \( f = 1 \ (m = 1) \) if the father (mother) has STEM-degree qualification, \( f = 2 \ (m = 2) \) if the father (mother) has a non-STEM degree qualification, \( f = 3 \ (m = 3) \) if the father (mother) has a high school (HS) qualification\textsuperscript{16} and \( f = 4 \ (m = 4) \) if the father (mother) has a junior high school (JHS) qualification or less.

\textsuperscript{15}We consider linear probability models to ease interpretation given the large number of interdependent effects we allow for in some of the specifications.

\textsuperscript{16}Our primary data source lacks information on field of study at high school (HS) for parents.
We experimented with different set of control variables in $X_i$. In the baseline regressions these controls include METTERE LISTA.

The key parameters of interest in equation (1) and (2) are $\alpha_f, \alpha_m$ and $\lambda_f, \lambda_m$, $f, m \in \{1, 2, 3, 4\}$: $\alpha_f, \alpha_m$ denote the effect of father or mother qualification on the probability that the student completes a STEM University degree, while $\lambda_f, \lambda_m$ denote the effect of father or mother qualification on the probability that the student completes a STEM high school degree. Positive estimates of $\alpha_1, \alpha_1$ denote intergenerational persistence of STEM education at University, and we interpret positive estimates of $\lambda_1, \lambda_1$ as evidence of intergenerational persistence of STEM education at high school. The sum $\alpha_1 + \alpha_2$ (or $\lambda_1 + \lambda_2$) represents the differential effect of having a graduated father (graduated mother) with respect to a parent with less than Junior High School (JHS) qualification on the probability of graduating from a STEM degree at University. Similarly, $\lambda_1 + \lambda_2$ (or $\lambda_1 + \lambda_2$) represents the corresponding effect of completing a STEM high school.

Interactions among parents within the household might not be trivial, thus we also consider alternative specifications where our key regressors identify all possible combinations of parental education levels. We define $PEdu_{fm}$, $f, m \in \{1, 2, 3, 4\}$, a set of mutually exclusive dummy variables taking the value 1 if the qualification of the father is $f$ and the qualification of the mother is $m$. For instance, $PEdu_{11}$ takes the value 1 if both parents hold a STEM University degree qualification, and 0 otherwise; while $PEdu_{12}$ takes the value 1 if the father holds a STEM University degree qualification and the mother holds a non-STEM University degree qualification, and 0 otherwise. Specifically we estimate also equations (3) and (4), where $\beta_{fm}$ and $\psi_{fm}$ denote the differential effect of having a father with qualification $f$ and a mother with qualification $m$ on the probability of completing a STEM University degree or a STEM high-school degree, respectively, with respect to the case in which both parents hold a qualification equal or lower to a junior high school degree.

\[ STEM_{i2} = \beta_0 + \sum_{m,f} \beta_{mf} PEdu_{mf} + \beta_X X + \zeta_{i2} \]  
\[ STEM_{i1} = \psi_0 + \sum_{m,f} \psi_{mf} PEdu_{mf} + \psi_X X + \zeta_{i1} \]
3.1 Exploiting a panel dimension

It is well known that with a time-invariant regressor -like parental qualifications in our case- one cannot exploit repeated observations on the same subject to identify the impact net of individual unobserved heterogeneity. We can however rely on this fixed effect strategy to identify how the impact of parental qualifications varies over time. More specifically, we benefit of the availability of time-varying outcomes to estimate the following equation with fixed effects

\[
STEM_{it} = \sum_f \Delta \gamma_{f,.} FED_{i,f} 1(t = 2) + \sum_m \Delta \gamma_{m,.} MED_{i,m} + \Delta \gamma'_X X_{i1} (t = 2) + \alpha_i + \varepsilon_{it} \tag{5}
\]

where \(1(t = 2)\) denotes a dummy variable taking the value one in period 2, and 0 otherwise. Since in our case, we have a balanced sample with two observations per subject (\(T = 2\)) and all the key regressors are time-invariant, the specification above is equivalent to estimate equation (6)

\[
\Delta STEM_{i2} = \sum_f \Delta \gamma_{f,.} FED_{i,f} + \sum_m \Delta \gamma_{m,.} MED_{i,m} + \Delta \gamma'_X X_{i1} + \Delta \nu_{i,2} \tag{6}
\]

where \(\Delta \gamma_{f,.} = \alpha_f - \lambda_f\) for the effects of fathers’ qualifications and similarly for the effect of mothers’ qualifications \(\Delta \gamma_{m,.}\). Indeed, one can view the specification in equation (6) as a difference-in-differences estimator.\(^{17}\)

In short, we exploit the longitudinal dimension of the data, specifically the fact that we observe the same student first completing high school, and later completing University, to control for time-invariant unobserved factors that may be related to parental qualifications and at the same time affect student qualifications. Our regressors of interest -namely the indicators for parental qualifications- are time-invariant. They proxy parental role models,\(^{17}\)

\[^{17}\text{In our setting, with time-invariant regressors, in large samples:}\]

\[
\hat{\Delta \gamma}_{f,.}^{FE} \approx \hat{\alpha}_{f,.}^{OLS} - \hat{\lambda}_{f,.}^{OLS} \tag{7}
\]

In other words, the fixed-effect estimator on a balanced panel with time invariant regressors is equivalent to the difference-in-differences estimator in equation (6), i.e. to the difference in the OLS estimators at the two stages of education in equation (1) and in equation (2). All these estimators for the difference in the impact of intergenerational persistence at the two stages of education -namely high school and university- are not affected by time-invariant individual unobserved heterogeneity.
rather than parental investments in their offspring, which are, conversely, typically assumed to be time-varying (Del Boca et al., 2017). To illustrate how the longitudinal dimension of our data is used in the empirical analysis consider for simplicity the ideal situation in which we could observe a time-varying indicator of parental investment of mothers \( MI_{it} \) and fathers \( FI_{it} \) on child \( i \) at time \( t \). A flexible way to model the relationship between the probability of graduating from a STEM degree at time \( t \) is presented in equation (8) below, where we allow for the effect of parental investment to be time-varying and we omit exogenous control variates \( X \) without loss of generality.

\[
STEM_{it} = \tau_{ft}FI_{it} + \tau_{mt}MI_{it} + \mu_i + \vartheta_{it} \tag{8}
\]

where \( \mu_i \) denotes a student-individual fixed effect and \( \vartheta_{it} \) represents an idiosyncratic error term. We observe each student \( i \) in two periods: at University \( t = 2 \) and at high school \( t = 1 \). By taking first differences of equation (9) we get

\[
\Delta STEM_{i2} = (\tau_{f2} - \tau_{f1})FI_{i2} + \tau_{f1}(FI_{i2} - FI_{i1}) + (\tau_{m2} - \tau_{m1})MI_{i2} + \tau_{m1}(MI_{i2} - MI_{i1}) + \Delta \vartheta_{i2} \tag{9}
\]

Equation 9 clearly illustrates that in this context the variability over time in the outcome \( STEM_{it} \) can be attributed to: (i) the change in the impact of a given parental investment over time, e.g. the variation \( (\tau_{f2} - \tau_{f1}) \) in the case of fathers; and (ii) the change in the amount of parental investment over time, for instance for fathers the variation \( (FI_{i2} - FI_{i1}) \).

With time-varying parental investments measures one could in principle identify and estimate both the variation in the effect of a given parental input in time 2 and time 1 \( ((\tau_{f2} - \tau_{f1}) \) for fathers and \( (\tau_{m2} - \tau_{m1}) \) for mothers) and the effect of a given parental input at time 1 \( (\tau_{f1} \) and \( \tau_{m1} \) for mothers and fathers, respectively).

However, in our context we do not have time-varying measures of parental investments but rather parental role models, i.e. a situation in which the regressor of interest is de facto time-invariant i.e. as if \( FI_{i2} \equiv FI_{i1} \) and \( MI_{i2} \equiv MI_{i1} \). Incorporating this restriction in equation (9), it becomes clear that the fixed effect strategy illustrated above allows to identify whether there is a differential effect in the second period of parental role models with respect to the effect of the same regressor in the first period, but not the absolute magnitude of these effects. In our context, the fixed effect identification strategy is equivalent to a difference-
in-differences strategy to identify this differential effect because there is no time-varying unobserved factor threatening the internal validity in the balanced panel with time-invariant regressors.

The strategies illustrated in this subsection allow to identify if the effects of parental qualifications – and specifically the field of parental education – are time-varying. In other words, it allows us to address the question of whether the effects of parental field of study are stronger at University or at high school. At the same time, it will not inform about the absolute magnitude of either of these effects.

One can also decompose the effects that parental qualification at University and at high school is to include an indicator of the students’ choice at high school as control in equation (1). We discuss this approach in the next section.

### 3.2 Direct and indirect effects of parental field of education

To disentangle the effects that parental field of education exert at high school – and then propagate to the choice of the University degree and its completion) – from those exerted directly on University graduation we also estimate equation (10)

\[
STEM_{i2} = \delta_0 + \sum_f \delta_f.FEd_{uf} + \sum_m \delta_m.MEd_{um} + \delta_{HS,STEM}STEM_{i1} + \delta_X'X_i + \omega_i
\]

where \(\delta_f, \delta_m\ \forall m, f \in \{1, 2, 3, 4\}\) identify - with some abuse of notation - the “direct” effect of father and mother education, respectively, on the choice of a STEM degree at University. Indeed, one can relate the parameters of equations (1), (2) and (10) as follows \(\alpha_f = \delta_{HS,stem}\lambda_f + \delta_f\ \forall f \in \{1, 2, 3, 4\}\), for father education. The same can be done for mother education.\(^{18}\) The so-called total effect of father qualification \(\alpha_f\) on the probability that the child completes a University degree in STEM is the sum of the direct effect exerted on University completion \(\delta_f\) and the indirect effect exerted through parental influence on a STEM high school completion \(\delta_{HS,stem}\lambda_f\). If the true value \(\delta_{HS,stem}\) were zero, there would be no linear dependence between the completion of STEM degree at high school and a STEM degree at University and the possibility of indirect effects of parents through this

\(^{18}\)It follows also that \(\Delta\gamma_f = (\delta_{HS,stem}\lambda_f + \delta_f) - \lambda_f\ \forall f \in \{1, 2, 3, 4\}\), for instance for father education, where \(\Delta\gamma_f\) are the parameters in equation (5). The same can be done for mother education.
A specific channel would be ruled out. When $\delta_{HS, stem} \neq 0$, i.e. if we can rule out the hypothesis that the two events are uncorrelated, the relative magnitude of the indirect effect through this specific channel depends on the size of $\lambda_f$, namely the intergenerational transmission of STEM field qualifications at high school.

In addition to exploring the sensitivity of the estimates of equation (1) to the inclusion of STEM high school completion, we experiment the sensitivity of our estimates to a larger set of controls (see section 5 below).

4 Evidence on intergenerational transmission of STEM field qualifications

The OLS estimates of the coefficients of the STEM Degree equation (1) are displayed in Table 7, while Table 8 presents estimates of equation (3), where interdependent effects between parents are allowed for. Table 9 and Table 10 presents the OLS estimates of the coefficients of the STEM High School (HS) equations with independent and interdependent parents’ effect respectively (corresponding to equations (2) and (4)).

Inspection of Table 7 reveals several interesting patterns. The effects of parents’ qualification on the probability of achieving a STEM University degree appears as a sizeable one. With respect to their counterparts with parents without an high school degree, students with graduated parents - regardless the field - are significantly more likely to achieve a STEM degree. Graduated fathers bring an advantage of about 14 p.p., and the figure is only slightly lower, about 12 p.p., for the mother.

The role of fathers and mothers in determining the University outcome of their offsprings become more differentiated when their field of study is taken into account. Having a father with a STEM degree makes about 16 p.p. for his child to achieve a STEM field degree - a magnitude that would almost counterbalance the observed gender gap in STEM degree completion in our sample - while the corresponding figure for STEM mothers is 10 p.p., i.e a substantially lower. STEM educated parents appear equally important for daughter and sons, as testified by the insignificant gender gap across student’s gender.\(^\text{19}\)

\(^\text{19}\)However, NON-STEM graduated fathers influence negatively the probability that their sons, but not
The comparison of Table 7 with Table 9 shows that a different scenario characterizes the previous educational stage. Indeed, STEM high school outcomes appear similarly affected by fathers and mothers qualifications and field of study. Moreover, in this adolescence period females are documented to be generally more sensitive than males to parental influence: all gender gaps are estimated to be negative, denoting that girls react more to both father’s and mother’s example, indicating the absence of a same-sex relationship.

In Table 8 the intergenerational coefficient of each parent is allowed to vary across educational level of the other parent. Scrolling through the first column, a very asymmetric behaviour in the completion of STEM University degree is observed in response to father and mother level and field of education: with respect to a student whose parents both hold a STEM degree, the probability that a student graduates in STEM at University decreases of about 7 (≈ 26 - 19) p.p if the mother holds a NON-STEM degree, while it decreases of about 20 (≈ 26 - 6) p.p. if it is the father that holds a NON-STEM degree. In other words, the influence of STEM fathers seems to prevail on the influence of STEM mothers. Moreover, the prominent role of fathers occurs for both sons and daughters, while STEM mothers’ role model reaches mainly daughters.

Table 10 confirms the more equal role of both parents for high school achievements, with STEM mothers stimulating STEM studies of their offsprings even in the presence of a NON-STEM father. The negative and statistically significant student gender gaps (in all cases but those in which one parent is low educated) describe girls as more affected by parental influence than their males counterparts in this phase of their life.

4.1 Fixed Effects Estimates

Table 11 displays the fixed effects coefficients corresponding to equation (5), which identify -net of time constant unobserved heterogeneity- the change in intergenerational transmission across the two educational stages we observe. These results point to effects that are gendered along both the parent’s and student’s dimensions. Indeed, the role of STEM fathers in determining the student’s STEM outcome is stronger for higher education than for high their daughters, achieve a STEM degree, while NON-STEM graduated mothers seem to encourage their daughters, but not their sons, to follow STEM tracks.
school (the effect significantly increases of approximately 4 p.p.), while STEM mothers matter more when their children are younger (the effect significantly decreases of $\approx 2$ p.p.). This is coherent with the view that fathers take a prominent role for children as they approach their entrance into the job market.

Moreover, as it can be observed in the last column, the gender difference in the time variation of the intergenerational coefficient is significant for both parents and it moves according to a same-gender pattern. Across time, STEM fathers become more influential for the STEM outcomes for their sons that for their daughters, while STEM mothers lose the role model they exerted for their daughters.

Table 11 conveys the same type of evidence allowing for interdependent effects, that show again -similar to their OLS counterparts in previous Table 8- the prevailing weight of STEM fathers with respect to STEM mothers.

### 4.2 Direct and Indirect Effects Estimates

In this section, we present estimates of equation (10), through which we decompose the effects that parental qualification exerts at University and at high school. Results are reported in Table 13 and Table 14.\(^{20}\) Consistently with descriptive statistics reported in Table 5, $\delta_{HS,STEM}$ is statistically significant: students who complete a STEM field high school are generally more likely (about 25 p.p.) to complete a STEM field degree at University. The corresponding non-parametric estimates resulting from the raw data in Table 5 is 29 p.p. ($=64%-35\%$). This is only proportional to the intergenerational transmission of parental qualification (see section 3.2): for instance, it would imply about a 3 p.p. $(0.25\cdot0.13)$ increase in the probability of completing a STEM University degree for students who completed a STEM high school and have a STEM graduated parent. This figure coincides with the estimates of the effect of attending at high school a curriculum that implies five more hours of science (per week, for one year) on the probability of graduating in a STEM field obtained by De Philippis (2017) for high ability 14-years-old students in England. In De Philippis (2017) these additional five hours correspond to a non-compulsory advanced science module.

\(^{20}\)In Table 13 we augment the specification in equation (1) with the control for the field choice of the student at high school, while in Table 14 we add the same control to the specification in equation (3).
According to the Italian regulation (Decreto Legislativo N. 226 17/10/2005), for instance the curriculum of students enrolled in a *liceo classico* (NON-STEM, 5 years curriculum) must have at least 858 hours of STEM (363 hours of Mathematics, 198 hours of Physics, 297 hours of Science), while the curriculum of students enrolled in a *liceo scientifico* (STEM, 5 years curriculum) must have at least 1485 hours of STEM (627 Mathematics, 429 of Physics, 429 of Science). Thus, in this example, the difference of STEM field education between the two curricula amounts to 627 hours in five years, about 125 hours per year (about 35 weeks of school), roughly 4 hours per week for five years. Building on the estimates of De Philippis (2017) for England, this should lead to about a 15% increase in the probability of graduating in a STEM field, an estimate about 40% lower that the 25 p.p. we document. This might be due to both differences in the Italian and British educational system and to the fact that in Italy, also within STEM high school and NON-STEM high school there is some heterogeneity in the number of hours devoted to STEM. In other words, the difference in STEM hours might be even larger than the one mentioned above between the curriculum of the *liceo classico* and *liceo scientifico*: indeed, in this example we could not include hours of Technology. Gender differences emerge as the persistence in educational choices is higher for males than for females: the effect of completing a STEM high school degree is 12 p.p. higher for males than for females. These conclusions are unaffected by the way in which we control for parental qualifications, i.e. they hold both referring to Table 13 and to Table 14 results.

The effect of parental qualifications is however not totally mediated by what happens at high school: the “direct” effect of father and mother education on the choice of a STEM degree at University is reflected in the sign and magnitude of the estimates of coefficients $\delta_f, \delta_m \forall m, f \in \{1, 2, 3, 4\}$, respectively (see Table 13). Parents’ qualification exert a still sizeable role: with respect to their counterparts with parents without at least a junior high school degree, students with graduate fathers (mothers) are about 11 (9) p.p. more likely to graduate in a STEM field at University (see the estimates of $\delta_1 + \delta_2$ and $\delta_1 + \delta_2$ in Table 13). These effects are statistically significant and only 20% lower for fathers (25% lower for mothers) than the corresponding estimates in Table 7, where we do not control for STEM high school completion. Similarly to what we observe in Table 7, we find that parental field
of study matters: STEM father graduates lead to 12 p.p. increase in the probability of graduating in STEM fields, compared to a substantially lower increase of 7 p.p. in the case of STEM graduate mothers. Consistently with previous findings, the gender gap (see column (4) in Table 13) is generally not significant: daughters and sons are equally influenced by STEM educated parents at this stage, even controlling for the mediating effect of high school completion. NON-STEM graduate fathers instead negatively influence their sons but not their daughters, while non-graduate mothers encourage their daughters to complete a STEM University degree but not their sons.

Allowing for the mediating role of high school, the asymmetric behaviour in response to father and mother level and field of education detected in Table 8 is confirmed by the estimates reported in Table 14: with respect to a student whose parents both hold a STEM degree, the probability that the student completes a STEM University degree decreases by 5 (≈ 20-15) p.p. if the mother holds a NON-STEM degree, while it decreases by 20 (≈ 20-0) p.p. if the same holds for the father. Consistently with the equal role of both parents at high school documented in Table 10, even conditioning on STEM high school completion, fathers’ influence on STEM University completion dominates the one of mothers.

Figures 1, 2, 3 offer a visual representation of the relative size of total, direct and indirect effects described above in the pooled sample, in the sample of male and in the sample of female students, respectively. It is apparent that the direct effect exert a the major role for STEM graduate parents but this is not true for NON-STEM graduates or parents with lower qualifications.

5 Further Robustness and heterogeneity analysis

5.1 On the role of parental (liberal) professions

Aina and Nicoletti (2018) study the intergenerational transmission of liberal professions and find that fathers exert a prominent role. The same authors also document an important role of intergenerational transmission of formal education on some of the compulsory steps required to become a liberal professional. Prompted by this intriguing results in the literature and the
consistency with our own findings on the intergenerational transmission of STEM education (see Section 4), we explore whether our results are driven by intergenerational transmission of liberal professions. To do so, we augment our regression with the interaction of parental qualifications with a dummy variable capturing whether the parent is a liberal professional (see equation (1)). Results are reported in Table 15. With some abuse of notation, we use \( \alpha_f, \alpha_m \) to denote the effects of parental qualifications, as in the baseline. The coefficients \( \alpha_{L1}, \alpha_{L2}, \alpha_{L3}, \alpha_{L4} \) denote the differential influence on the probability of graduating in a STEM field of father’s qualifications when the father is a liberal professional; similarly \( \alpha_{L1}, \alpha_{L2}, \alpha_{L3}, \alpha_{L4} \) denote the differential influence of mother’s qualifications when the mother is a liberal professional. These differential influence is statistically significant for STEM graduate liberal professional fathers and amounts to about a 5 p.p. increase with respect to the positive intergenerational transmission of STEM education between fathers and children (14 p.p.). Our baseline results are confirmed, as the intergenerational transmission of STEM field education at University \( (\alpha_1) \) is only 12% lower than in Table 7. For STEM graduate mothers no statistically significant differential effect of holding a liberal profession can be detected. NON-STEM graduate liberal professional fathers tend to reduce the probability that children graduate in STEM. This effect is 4 p.p. larger in absolute terms with respect to the effect for NON-STEM graduate fathers who are not liberal professional. While non-stem graduate mothers who are not liberal professional tend to positively affect the probability that children graduate in STEM, these positive effect vanishes for NON-STEM graduate mothers who are liberal professional \( (= 0.0202 - 0.0172 = 0.003) \). Notably, the result in the pooled sample is only driven by daughters for non-liberal professional mothers and by sons for liberal professional mothers, with statistically significant differences along the gender of the child.

In short, we now can document a significant differential role of fathers’ occupation: liberal professional STEM graduate fathers increase the likelihood of children graduating from a STEM field, for both males and females, while liberal professionals non-STEM graduate fathers reduce it, more for sons than for daughters. Liberal professional mothers exert

\[ \text{All regression presented so far already include the following set of controls: region of residence, social class and parents’ jobs, i.e. a set of dummy variables including a binary indicator for self-reported liberal professional. We now interact the latter with the educational dummies.} \]
negligible roles for daughters and tend to reduce the probability of graduating from a STEM field for sons. While the STEM-field intergenerational persistence at University is not related to having parents who are liberal professional, we find that the influence of NON-STEM graduate fathers is mainly driven by those who are liberal professional. If entry barriers into liberal professions are higher in NON-STEM liberal jobs (e.g. notary or lawyer) compared to STEM liberal jobs (e.g. engineer), this finding would be consistent with the discussion by (Aina and Nicoletti, 2018, Tab.1, pg.111), who suggest that “high entry barriers into the profession increase the occupational transmission from fathers to children” and find that “non-graduate liberal professionals transmit to their child a level of formal human capital similar to the blue-collar’s one and lower than the entrepreneurs’ one” (see pg.115). Our results are consistent with the view that the intergenerational transmission of the NON-STEM liberal profession happens more for sons than for daughters: in Table 15 $\hat{\alpha}_{L,2}$ (NON-STEM graduate liberal professional fathers) is -0.09 for males and -0.02 for females; $\hat{\alpha}_{L,2}$ (NON-STEM graduate liberal professional mothers) is -0.05 for males and 0 for females, and all non-zero estimates are statistically significant at 5%.

In short, our intergenerational correlations of STEM qualifications are not driven by the transmission of liberal profession.

5.2 On the role of student observables

We perform further robustness checks by controlling for student’s characteristics observable at University (but not at high school) in order to document whether additional controls will change our key parameters of interest, i.e. the effect of father and mother qualification on the probability that the students completes a STEM degree at University. We increment the set of control variables included in the baseline model defined in equation (1) and (3) by incorporating also the location of the University with respect to the region of residence, the type of high school completed, academic and work values reported by students as being important in their degree choice at University. The OLS estimates of the coefficients of equation (1) with enhanced controls are presented in Table Online Appendix - 9 while Table Online Appendix - 10 displays estimates of equation (3), respectively. Given that in this model specification we include controls for the type of high school, the coefficients $\alpha_f, \alpha_m$
presented in Table Online Appendix - 9 (Table Online Appendix - 10) should be compared with coefficients $\delta_f, \delta_m (\theta_{fm})$ of equation (10) reported in Table 13 (Table 14): indeed, also in this case, they represent the so called “direct” effect of father and mother education on the choice of a STEM degree at University. From the tables, we can see that the effect of father’s and mother’s qualification on the probability that the child graduates in STEM at University remain mostly stable even with the insertion of additional controls. Table Online Appendix - 11 and Table Online Appendix - 12 display the estimates for parental education impact when adding also high school final mark to the list of controls. As before, the estimates of the intergenerational correlation of parents’ and child’s education do not vary considerably, i.e., we notice variations of around 1 p.p. (about 8% of the estimates of “direct” effect of father STEM education and 6% of the estimates of the corresponding total effect reported in Table 7) for some of the coefficients, while other remain unchanged.

6 Concluding remarks

We document the presence of sizeable intergenerational transmission of STEM education in Italy at both the high-school and University level. In our sample, 70% of male students complete a STEM high school, while the corresponding figure for girls is 40%. The gender gap is significantly smaller at University; 34% of males graduate from a STEM field at University compared to 15% of females. Indeed, the field choice is not persistent as one would expect: a substantial fraction of students graduating from a STEM degree at University does not have a high school STEM degree (21% in the pooled sample, 33% among females). We find that having a father who graduated from a STEM field increases the likelihood of graduating from a STEM field at high school by 13 p.p. (9 p.p. for sons; 12 p.p for daughters) and at University by 16 p.p. (16 p.p. for sons; 14 for daughters). The magnitude of these figures is half of the gender gap in STEM degree completion at high school and as large as the gap in STEM University degree completion. The extent of intergenerational trasmission is sizeable also for mothers: having a STEM graduate mother increases the probability of completing a STEM high school by 12 p.p. (7 p.p. for sons; 14 p.p. for daughters) and the probability of graduating from a STEM field at University by 10 p.p. (9 p.p. for sons; 10
p.p. for daughters). Exploiting the non-negligible variability in STEM field between high-
school and University graduation rates, we can document that, controlling for time-invariant
unobservables, the intergenerational transmission is stronger at University. While at high
school parents exert a similar role, at University the influence of STEM fathers generally
prevails on the influence of STEM mothers. The prominent role of fathers reaches both sons
and daughters, while STEM mothers’ role model reaches mainly daughters.\textsuperscript{22}

Our findings on the persistence of STEM educational attainments are consistent with
previous empirical evidence on intergenerational transmission of education in Italy (Checchi
et al., 2013) and we uncover interesting novel patterns that could inform parents and policy
makers concerned with the shortage of supply of STEM related skills, particularly among
females.

\textsuperscript{22}We do not have measures of active parental investments on their offspring and interpret our measures
of parental educational qualifications as proxies of time-invariant role models.
References


Brenoe, A., 2017. Sibling Gender Composition and Participation in STEM Education. mimeo, University of Copenhagen, Department of Economics.


Granato, S., 2018. Early Influences and the Gender Gap in STEM. mimeo, Queen Mary University of London.


Winters, J., 2015. Do higher levels of education and skills in an area benefit wider society? IZA World of Labor.


### Tables and Figures

#### Table 1: Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>female*</td>
<td>0.605</td>
<td>0.489</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>STEM_{i2} (Degree)*</td>
<td>0.231</td>
<td>0.422</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>STEM_{i1} (HS)</td>
<td>0.518</td>
<td>0.500</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Father education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEdu_{11} (STEM degree)</td>
<td>0.062</td>
<td>0.242</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>FEdu_{12} (non-STEM degree)</td>
<td>0.143</td>
<td>0.351</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>FEdu_{31} (HS)</td>
<td>0.456</td>
<td>0.498</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>FEdu_{41} (JHS or less)</td>
<td>0.338</td>
<td>0.473</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Mother education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEdu_{11} (STEM degree)</td>
<td>0.035</td>
<td>0.185</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MEdu_{12} (non-STEM degree)</td>
<td>0.156</td>
<td>0.362</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MEdu_{31} (HS)</td>
<td>0.505</td>
<td>0.500</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MEdu_{41} (JHS or less)</td>
<td>0.304</td>
<td>0.460</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Parents’ combined education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEdu_{11}</td>
<td>0.011</td>
<td>0.106</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{12}</td>
<td>0.022</td>
<td>0.147</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{13}</td>
<td>0.026</td>
<td>0.158</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{14}</td>
<td>0.003</td>
<td>0.055</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{21}</td>
<td>0.011</td>
<td>0.105</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{22}</td>
<td>0.066</td>
<td>0.247</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{23}</td>
<td>0.058</td>
<td>0.234</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{24}</td>
<td>0.008</td>
<td>0.092</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{31}</td>
<td>0.011</td>
<td>0.104</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{32}</td>
<td>0.056</td>
<td>0.230</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{33}</td>
<td>0.291</td>
<td>0.454</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{34}</td>
<td>0.097</td>
<td>0.297</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{41}</td>
<td>0.002</td>
<td>0.045</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{42}</td>
<td>0.012</td>
<td>0.107</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{43}</td>
<td>0.130</td>
<td>0.336</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PEdu_{44}</td>
<td>0.195</td>
<td>0.396</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Observations</td>
<td>155,603</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Notation: $STEM_{i2}$ takes the value 1 if the individual graduates from a STEM field degree at University; $STEM_{i1}$ takes the value 1 if the individual graduates from a STEM field degree at high school; $FEdu_{i1}$, $MEdu_{j}$, $PEdu_{ij}$, $i,j \in \{1,2,3,4\}$ are dummy variables denoting levels of parental education as clarified in the table. Variables marked with * indicate administrative data, unmarked variables indicate survey data.
Table 2: Association between parents’ education and child’s choices

<table>
<thead>
<tr>
<th></th>
<th>STEM</th>
<th>non-STEM</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. %</td>
<td>No. %</td>
<td>No. %</td>
</tr>
<tr>
<td><strong>Father’s education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_{Edu_1}$ (STEM degree)</td>
<td>3,716 (38.31)</td>
<td>5,984 (61.69)</td>
<td>9,700 (100)</td>
</tr>
<tr>
<td>$F_{Edu_2}$ (non-STEM degree)</td>
<td>4,353 (19.50)</td>
<td>17,974 (80.50)</td>
<td>22,327 (100)</td>
</tr>
<tr>
<td>$F_{Edu_3}$ (HS)</td>
<td>16,953 (23.90)</td>
<td>53,976 (76.10)</td>
<td>70,929 (100)</td>
</tr>
<tr>
<td>$F_{Edu_4}$ (JHS or less)</td>
<td>10,977 (20.85)</td>
<td>41,670 (79.15)</td>
<td>52,647 (100)</td>
</tr>
<tr>
<td><strong>Mother’s education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_{Edu_1}$ (STEM degree)</td>
<td>1,911 (34.65)</td>
<td>3,604 (65.35)</td>
<td>5,515 (100)</td>
</tr>
<tr>
<td>$M_{Edu_2}$ (non-STEM degree)</td>
<td>5,694 (23.51)</td>
<td>18,523 (76.49)</td>
<td>24,217 (100)</td>
</tr>
<tr>
<td>$M_{Edu_3}$ (HS)</td>
<td>18,448 (23.46)</td>
<td>60,185 (76.54)</td>
<td>78,633 (100)</td>
</tr>
<tr>
<td>$M_{Edu_4}$ (JHS or less)</td>
<td>9,946 (21.06)</td>
<td>37,292 (78.94)</td>
<td>47,238 (100)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>35,999 (23.14)</td>
<td>119,604 (76.86)</td>
<td>155,603 (100)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis.

Table 3: Association between parents’ education and child’s choices by gender: males

<table>
<thead>
<tr>
<th></th>
<th>STEM</th>
<th>non-STEM</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. %</td>
<td>No. %</td>
<td>No. %</td>
</tr>
<tr>
<td><strong>Father’s education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_{Edu_1}$ (STEM degree)</td>
<td>2,153 (48.51)</td>
<td>2,285 (51.49)</td>
<td>4,438 (100)</td>
</tr>
<tr>
<td>$F_{Edu_2}$ (non-STEM degree)</td>
<td>2,568 (25.31)</td>
<td>7,578 (74.69)</td>
<td>10,146 (100)</td>
</tr>
<tr>
<td>$F_{Edu_3}$ (HS)</td>
<td>10,274 (35.76)</td>
<td>18,454 (64.24)</td>
<td>28,728 (100)</td>
</tr>
<tr>
<td>$F_{Edu_4}$ (JHS or less)</td>
<td>6,402 (35.33)</td>
<td>11,720 (64.67)</td>
<td>18,122 (100)</td>
</tr>
<tr>
<td><strong>Mother’s education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_{Edu_1}$ (STEM degree)</td>
<td>1,096 (43.77)</td>
<td>1,408 (56.23)</td>
<td>2,504 (100)</td>
</tr>
<tr>
<td>$M_{Edu_2}$ (non-STEM degree)</td>
<td>3,380 (31.42)</td>
<td>7,379 (68.58)</td>
<td>10,759 (100)</td>
</tr>
<tr>
<td>$M_{Edu_3}$ (HS)</td>
<td>11,046 (34.73)</td>
<td>20,762 (65.27)</td>
<td>31,808 (100)</td>
</tr>
<tr>
<td>$M_{Edu_4}$ (JHS or less)</td>
<td>5,875 (35.90)</td>
<td>10,488 (64.10)</td>
<td>16,363 (100)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>21,397 (34.83)</td>
<td>40,037 (65.17)</td>
<td>61,434 (100)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis.
### Table 4: Association between parent’s education and child’s choices by gender: females

<table>
<thead>
<tr>
<th></th>
<th>STEM</th>
<th>non-STEM</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
<td>No.</td>
</tr>
<tr>
<td><strong>Father’s education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F Edu_1$ (STEM degree)</td>
<td>1,563 (29.70)</td>
<td>3,699 (70.30)</td>
<td>5,262 (100)</td>
</tr>
<tr>
<td>$F Edu_2$ (non-STEM degree)</td>
<td>1,785 (14.65)</td>
<td>10,396 (85.35)</td>
<td>12,181 (100)</td>
</tr>
<tr>
<td>$F Edu_3$ (HS)</td>
<td>6,679 (15.83)</td>
<td>35,522 (84.17)</td>
<td>42,201 (100)</td>
</tr>
<tr>
<td>$F Edu_4$ (JHS or less)</td>
<td>4,575 (13.25)</td>
<td>29,950 (86.75)</td>
<td>34,525 (100)</td>
</tr>
<tr>
<td><strong>Mother’s education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M Edu_1$ (STEM degree)</td>
<td>815 (27.07)</td>
<td>2,196 (72.93)</td>
<td>3,011 (100)</td>
</tr>
<tr>
<td>$M Edu_2$ (non-STEM degree)</td>
<td>2,314 (17.19)</td>
<td>11,144 (82.81)</td>
<td>13,458 (100)</td>
</tr>
<tr>
<td>$M Edu_3$ (HS)</td>
<td>7,402 (15.81)</td>
<td>39,423 (84.19)</td>
<td>46,825 (100)</td>
</tr>
<tr>
<td>$M Edu_4$ (JHS and less)</td>
<td>4,071 (13.19)</td>
<td>26,804 (86.81)</td>
<td>30,875 (100)</td>
</tr>
</tbody>
</table>

Pearson $\chi^2 = 953.5836 \ Pr=0.000$

**Total** | 14,602 (15.51) | 79,567 (84.49) | 94,169 (100) |

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis.

### Table 5: Persistence of child’s field choices over time (at high schools versus at university): frequency, (row percentage), [column percentage], {cell percentage}. Pooled, females

<table>
<thead>
<tr>
<th>STEM field at high school $STEM_{i1}$</th>
<th>STEM field at University $STEM_{i2}$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>0</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>67,396 (51.67)</td>
<td>7,505 (4.83)</td>
</tr>
<tr>
<td></td>
<td>(89.98) (91.44)</td>
<td>(10.02) (8.56)</td>
</tr>
<tr>
<td></td>
<td>[56.35] [64.94]</td>
<td>[20.85] [33.11]</td>
</tr>
<tr>
<td></td>
<td>{43.31} {54.87}</td>
<td>{4.82} {5.13}</td>
</tr>
<tr>
<td><strong>1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>52,208 (27.89)</td>
<td>28,494 (9.76)</td>
</tr>
<tr>
<td></td>
<td>(64.69) (74.07)</td>
<td>(35.31) (25.93)</td>
</tr>
<tr>
<td></td>
<td>[43.65] [35.06]</td>
<td>[79.15] [66.89]</td>
</tr>
<tr>
<td></td>
<td>{33.55} {29.62}</td>
<td>{18.31} {10.37}</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>119,604 (79.567)</td>
<td>35,999 (14.602)</td>
</tr>
<tr>
<td></td>
<td>(76.86) (84.49)</td>
<td>(23.14) (15.51)</td>
</tr>
<tr>
<td></td>
<td>{76.86} {84.49}</td>
<td>{23.14} {15.51}</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis.
Table 6: Association between parental education levels and field choice: frequency, (row percentage), [column percentage], {cell percentage}

<table>
<thead>
<tr>
<th>Father’s education</th>
<th>JHS or less $M_{Edu.4}$</th>
<th>HS $M_{Edu.3}$</th>
<th>non-STEM degree $M_{Edu.2}$</th>
<th>STEM degree $M_{Edu.1}$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>JHS or less $F_{Edu.4}$</td>
<td>30,269 (57.49)</td>
<td>20,242 (38.45)</td>
<td>1,820 (3.46)</td>
<td>316 (0.60)</td>
<td>52,647 (100.00)</td>
</tr>
<tr>
<td></td>
<td>30,269 [64.08]</td>
<td>20,242 [25.74]</td>
<td>1,820 [7.52]</td>
<td>316 [5.73]</td>
<td>52,647 [33.85]</td>
</tr>
<tr>
<td></td>
<td>9.45 {13.01}</td>
<td>1.17 {1.17}</td>
<td>0.20 {0.20}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS $F_{Edu.3}$</td>
<td>15,170 (21.39)</td>
<td>45,327 (63.90)</td>
<td>8,739 (12.32)</td>
<td>1,693 (2.39)</td>
<td>70,929 (100.00)</td>
</tr>
<tr>
<td>non-STEM degree $F_{Edu.2}$</td>
<td>1,320 (5.91)</td>
<td>9,077 (40.65)</td>
<td>10,195 (45.66)</td>
<td>1,735 (7.77)</td>
<td>22,327 (100.00)</td>
</tr>
<tr>
<td>STEM degree $F_{Edu.1}$</td>
<td>479 (4.94)</td>
<td>3,987 (41.10)</td>
<td>3,463 (35.70)</td>
<td>1,771 (18.26)</td>
<td>9,700 (100.00)</td>
</tr>
<tr>
<td>Total</td>
<td>47,238 (30.36)</td>
<td>78,633 (50.53)</td>
<td>24,217 (15.56)</td>
<td>5,515 (3.54)</td>
<td>155,603 (100.00)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis.
Table 7: Estimates of the effects of parental education on STEM degree choice at the University, pooled and by gender of the student

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(mean: 0.231)</td>
<td>(mean: 0.348)</td>
<td>(mean: 0.155)</td>
<td></td>
</tr>
<tr>
<td><strong>Father education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.1598***</td>
<td>0.1590***</td>
<td>0.1407***</td>
<td>0.0184</td>
</tr>
<tr>
<td></td>
<td>(0.0060)</td>
<td>(0.0098)</td>
<td>(0.0074)</td>
<td>(0.0122)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>-0.0185***</td>
<td>-0.0605***</td>
<td>-0.0029</td>
<td>-0.0575***</td>
</tr>
<tr>
<td></td>
<td>(0.0043)</td>
<td>(0.0075)</td>
<td>(0.0049)</td>
<td>(0.0090)</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.0253***</td>
<td>0.0217***</td>
<td>0.0170***</td>
<td>0.0047</td>
</tr>
<tr>
<td></td>
<td>(0.0028)</td>
<td>(0.0052)</td>
<td>(0.0029)</td>
<td>(0.0060)</td>
</tr>
<tr>
<td><strong>Mother education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.1015***</td>
<td>0.0914***</td>
<td>0.0955***</td>
<td>-0.0041</td>
</tr>
<tr>
<td></td>
<td>(0.0074)</td>
<td>(0.0119)</td>
<td>(0.0090)</td>
<td>(0.0149)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.0196***</td>
<td>0.0013</td>
<td>0.0205***</td>
<td>-0.0192**</td>
</tr>
<tr>
<td></td>
<td>(0.0044)</td>
<td>(0.0077)</td>
<td>(0.0049)</td>
<td>(0.0091)</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.0137***</td>
<td>-0.0025</td>
<td>0.0150***</td>
<td>-0.0175***</td>
</tr>
<tr>
<td></td>
<td>(0.0028)</td>
<td>(0.0053)</td>
<td>(0.0030)</td>
<td>(0.0061)</td>
</tr>
<tr>
<td>$\alpha_1 + \alpha_2$</td>
<td>0.1413***</td>
<td>0.0986***</td>
<td>0.1378***</td>
<td>-0.0392*</td>
</tr>
<tr>
<td></td>
<td>(0.0088)</td>
<td>(0.0151)</td>
<td>(0.0103)</td>
<td>(0.0183)</td>
</tr>
<tr>
<td>$\alpha_1 + \alpha_2$</td>
<td>0.1211***</td>
<td>0.0927***</td>
<td>0.1160***</td>
<td>-0.0233</td>
</tr>
<tr>
<td></td>
<td>(0.0099)</td>
<td>(0.0168)</td>
<td>(0.0117)</td>
<td>(0.0204)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of OLS estimates of equation (1) in the paper, reported here for convenience. $STEM_i = \alpha_0 + \alpha_1 F Edu_1 + \alpha_2 F Edu_2 + \alpha_3 F Edu_3 + \alpha_1 M Edu_1 + \alpha_2 M Edu_2 + \alpha_3 M Edu_3 X_i + \varepsilon_i$, where $STEM_i$ denotes a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise; $F Edu_i, M Edu_j, i,j \in \{1,2,3,4\}$ are dummy variables denoting the qualification level of fathers and mothers respectively, where $i = 1 (j = 1)$ if the father (mother) has STEM-degree qualification, $i = 2 (j = 2)$ if the father (mother) has a non-STEM degree qualification, $i = 3 (j = 3)$ if the father (mother) has a high school qualification and $i = 4 (j = 4)$ if the father (mother) has a junior high school (JHS) qualification or less. $X$ includes the following covariates: region of residence, social class and parents’ jobs. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table 8: Estimates of the effects of parental education on STEM degree choice at the University, pooled and by gender of the student

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(mean: 0.231)</td>
<td>(mean: 0.348)</td>
<td>(mean: 0.155)</td>
<td></td>
</tr>
<tr>
<td><strong>Both parents with with degree in STEM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>0.2516***</td>
<td>0.2375***</td>
<td>0.2325***</td>
<td>0.0050</td>
</tr>
<tr>
<td></td>
<td>(0.0125)</td>
<td>(0.0190)</td>
<td>(0.0162)</td>
<td>(0.0250)</td>
</tr>
<tr>
<td>$\beta_{12}$</td>
<td>0.1846***</td>
<td>0.1685***</td>
<td>0.1642***</td>
<td>0.0043</td>
</tr>
<tr>
<td></td>
<td>(0.0092)</td>
<td>(0.0145)</td>
<td>(0.0114)</td>
<td>(0.0185)</td>
</tr>
<tr>
<td>$\beta_{13}$</td>
<td>0.1811***</td>
<td>0.1685***</td>
<td>0.1628***</td>
<td>0.0057</td>
</tr>
<tr>
<td></td>
<td>(0.0085)</td>
<td>(0.0137)</td>
<td>(0.0103)</td>
<td>(0.0171)</td>
</tr>
<tr>
<td>$\beta_{14}$</td>
<td>0.1474***</td>
<td>0.1496***</td>
<td>0.1092***</td>
<td>0.0404</td>
</tr>
<tr>
<td></td>
<td>(0.0219)</td>
<td>(0.0343)</td>
<td>(0.0263)</td>
<td>(0.0432)</td>
</tr>
<tr>
<td><strong>Father with degree in STEM different mother education levels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.0579***</td>
<td>0.0129</td>
<td>0.0627***</td>
<td>-0.0498**</td>
</tr>
<tr>
<td></td>
<td>(0.0111)</td>
<td>(0.0179)</td>
<td>(0.0137)</td>
<td>(0.0225)</td>
</tr>
<tr>
<td>$\beta_{31}$</td>
<td>0.1554***</td>
<td>0.1459***</td>
<td>0.1373***</td>
<td>0.0086</td>
</tr>
<tr>
<td></td>
<td>(0.0121)</td>
<td>(0.0196)</td>
<td>(0.0148)</td>
<td>(0.0245)</td>
</tr>
<tr>
<td>$\beta_{41}$</td>
<td>0.1869***</td>
<td>0.2027***</td>
<td>0.1631***</td>
<td>0.0396</td>
</tr>
<tr>
<td></td>
<td>(0.0275)</td>
<td>(0.0456)</td>
<td>(0.0325)</td>
<td>(0.0560)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of OLS estimates of equation (3) in the paper on the pooled sample, on the sample of male students and on the sample of females students, respectively. The outcome is $STEM_{i2}$ a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise. All equations include $X$, the following set of covariates: region of residence, social class and parents’ jobs. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. $\beta_{ij}$, $i,j \in \{1,2,3,4\}$ denote the differential effect of having a father with qualification $i$ and a mother with qualification $j$ on the probability of completing a STEM University degree, respectively, with respect to the case in which both parents hold a qualification equal or lower to a junior high school degree. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table 9: Estimates of the effects of parental education on STEM choice at high school, pooled and by gender of the student

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1) (mean: 0.518)</th>
<th>Males (2) (mean: 0.700)</th>
<th>Females (3) (mean: 0.400)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Father education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.1261***</td>
<td>0.0898***</td>
<td>0.1231***</td>
<td>-0.0333***</td>
</tr>
<tr>
<td></td>
<td>(0.0064)</td>
<td>(0.0086)</td>
<td>(0.0086)</td>
<td>(0.0122)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.0194***</td>
<td>-0.0232***</td>
<td>0.0217***</td>
<td>-0.0448***</td>
</tr>
<tr>
<td></td>
<td>(0.0052)</td>
<td>(0.0075)</td>
<td>(0.0067)</td>
<td>(0.0101)</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>0.0473***</td>
<td>0.0285***</td>
<td>0.0429***</td>
<td>-0.0144**</td>
</tr>
<tr>
<td></td>
<td>(0.0033)</td>
<td>(0.0050)</td>
<td>(0.0040)</td>
<td>(0.0064)</td>
</tr>
<tr>
<td><strong>Mother education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.1233***</td>
<td>0.0746***</td>
<td>0.1411***</td>
<td>-0.0665***</td>
</tr>
<tr>
<td></td>
<td>(0.0078)</td>
<td>(0.0105)</td>
<td>(0.0107)</td>
<td>(0.0150)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.0187***</td>
<td>-0.0107</td>
<td>0.0214***</td>
<td>-0.0321***</td>
</tr>
<tr>
<td></td>
<td>(0.0052)</td>
<td>(0.0075)</td>
<td>(0.0066)</td>
<td>(0.0100)</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>0.0375***</td>
<td>0.0174***</td>
<td>0.0372***</td>
<td>-0.0198***</td>
</tr>
<tr>
<td></td>
<td>(0.0034)</td>
<td>(0.0051)</td>
<td>(0.0041)</td>
<td>(0.0066)</td>
</tr>
<tr>
<td>$\lambda_1 + \lambda_2$</td>
<td>0.1455***</td>
<td>0.0667***</td>
<td>0.1448***</td>
<td>-0.0781***</td>
</tr>
<tr>
<td></td>
<td>(0.0100)</td>
<td>(0.0141)</td>
<td>(0.0131)</td>
<td>(0.0192)</td>
</tr>
<tr>
<td>$\lambda_1 + \lambda_2$</td>
<td>0.1420***</td>
<td>0.0639***</td>
<td>0.1625***</td>
<td>-0.0986***</td>
</tr>
<tr>
<td></td>
<td>(0.0110)</td>
<td>(0.0155)</td>
<td>(0.0146)</td>
<td>(0.0213)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of OLS estimates of equation (2) in the paper, reported here for convenience. $STEM_{i1} = \lambda_0 + \lambda_1 FEdu_1 + \lambda_2 FEdu_2 + \lambda_3 MEdu_1 + \lambda_4 MEdu_2 + \lambda_5 MEdu_3 + \lambda_6 X_i + \varepsilon_{i1}$, where $STEM_{i1}$ denotes a dummy taking the value 1 if student $i$ completes a STEM field high school degree and 0 otherwise; $FEdu_{i,j}$, $MEdu_{i,j}$, $i,j \in \{1,2,3,4\}$ are dummy variables denoting the qualification level of fathers and mothers respectively, where $i = 1$ ($j = 1$) if the father (mother) has STEM-degree qualification, $i = 2$ ($j = 2$) if the father (mother) has a non-STEM degree qualification, $i = 3$ ($j = 3$) if the father (mother) has a high school qualification and $i = 4$ ($j = 4$) if the father (mother) has a junior high school (JHS) qualification or less. $X$ includes the following covariates: region of residence, social class and parents’ jobs. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table 10: Estimates of the effects of parental education on STEM choice at high school, pooled and by gender of the student

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(mean: 0.518)</td>
<td>(mean: 0.700)</td>
<td>(mean: 0.400)</td>
<td></td>
</tr>
<tr>
<td><strong>Both parents with degree in STEM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{11}$</td>
<td>0.2194***</td>
<td>0.1282***</td>
<td>0.2458***</td>
<td>-0.1176***</td>
</tr>
<tr>
<td></td>
<td>(0.0121)</td>
<td>(0.0161)</td>
<td>(0.0172)</td>
<td>(0.0235)</td>
</tr>
<tr>
<td>$\psi_{12}$</td>
<td>0.1402***</td>
<td>0.0872***</td>
<td>0.1303***</td>
<td>-0.0431**</td>
</tr>
<tr>
<td></td>
<td>(0.0096)</td>
<td>(0.0130)</td>
<td>(0.0131)</td>
<td>(0.0185)</td>
</tr>
<tr>
<td>$\psi_{13}$</td>
<td>0.1854***</td>
<td>0.1298***</td>
<td>0.1849***</td>
<td>-0.0551***</td>
</tr>
<tr>
<td></td>
<td>(0.0087)</td>
<td>(0.0117)</td>
<td>(0.0119)</td>
<td>(0.0167)</td>
</tr>
<tr>
<td>$\psi_{14}$</td>
<td>0.1825***</td>
<td>0.1482***</td>
<td>0.1537***</td>
<td>-0.0055</td>
</tr>
<tr>
<td></td>
<td>(0.0223)</td>
<td>(0.0270)</td>
<td>(0.0312)</td>
<td>(0.0413)</td>
</tr>
<tr>
<td><strong>Father with degree in STEM different mother education levels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{21}$</td>
<td>0.1404***</td>
<td>0.0631***</td>
<td>0.1508***</td>
<td>-0.0877***</td>
</tr>
<tr>
<td></td>
<td>(0.0128)</td>
<td>(0.0174)</td>
<td>(0.0177)</td>
<td>(0.0248)</td>
</tr>
<tr>
<td>$\psi_{31}$</td>
<td>0.2057***</td>
<td>0.1420***</td>
<td>0.2147***</td>
<td>-0.0728***</td>
</tr>
<tr>
<td></td>
<td>(0.0123)</td>
<td>(0.0160)</td>
<td>(0.0170)</td>
<td>(0.0234)</td>
</tr>
<tr>
<td>$\psi_{41}$</td>
<td>0.1945***</td>
<td>0.1696***</td>
<td>0.1910***</td>
<td>-0.0214</td>
</tr>
<tr>
<td></td>
<td>(0.0272)</td>
<td>(0.0338)</td>
<td>(0.0362)</td>
<td>(0.0495)</td>
</tr>
</tbody>
</table>

*Notes*: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of OLS estimates of equation (4) in the paper on the pooled sample, on the sample of male students and on the sample of females students, respectively. The outcome is $STEM_{i1}$, a dummy taking the value 1 if student $i$ graduates from a STEM field at high school and 0 otherwise. All equations include $X$, the following set of covariates: region of residence, social class and parents' jobs. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. $\psi_{ij} \in \{1, 2, 3, 4\}$ denote the differential effect of having a father with qualification $i$ and a mother with qualification $j$ on the probability of completing a STEM high-school degree, respectively, with respect to the case in which both parents hold a qualification equal or lower to a junior high school degree. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table 11: Heterogeneity of the effects of parental education on STEM degree choice at the University wrt to the same choice at High School, pooled and by gender of the student. Fixed effects (FE) estimates.

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Father education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta \gamma_1$</td>
<td>0.0338***</td>
<td>0.0692***</td>
<td>0.0176*</td>
<td>0.0516***</td>
</tr>
<tr>
<td></td>
<td>(0.0075)</td>
<td>(0.0113)</td>
<td>(0.0101)</td>
<td>(0.0152)</td>
</tr>
<tr>
<td>$\Delta \gamma_2$</td>
<td>-0.0379***</td>
<td>-0.0373***</td>
<td>-0.0246***</td>
<td>-0.0127</td>
</tr>
<tr>
<td></td>
<td>(0.0058)</td>
<td>(0.0092)</td>
<td>(0.0075)</td>
<td>(0.0119)</td>
</tr>
<tr>
<td>$\Delta \gamma_3$</td>
<td>-0.0220***</td>
<td>-0.0068</td>
<td>-0.0258***</td>
<td>0.0191**</td>
</tr>
<tr>
<td></td>
<td>(0.0036)</td>
<td>(0.0061)</td>
<td>(0.0044)</td>
<td>(0.0075)</td>
</tr>
<tr>
<td><strong>Mother education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta \gamma_1$</td>
<td>-0.0218**</td>
<td>0.0168</td>
<td>-0.0456***</td>
<td>0.0624***</td>
</tr>
<tr>
<td></td>
<td>(0.0094)</td>
<td>(0.0142)</td>
<td>(0.0126)</td>
<td>(0.0190)</td>
</tr>
<tr>
<td>$\Delta \gamma_2$</td>
<td>0.0009</td>
<td>0.0120</td>
<td>-0.0009</td>
<td>0.0129</td>
</tr>
<tr>
<td></td>
<td>(0.0058)</td>
<td>(0.0091)</td>
<td>(0.0074)</td>
<td>(0.0117)</td>
</tr>
<tr>
<td>$\Delta \gamma_3$</td>
<td>-0.0238***</td>
<td>-0.0199***</td>
<td>-0.0222***</td>
<td>0.0022</td>
</tr>
<tr>
<td></td>
<td>(0.0037)</td>
<td>(0.0062)</td>
<td>(0.0045)</td>
<td>(0.0077)</td>
</tr>
<tr>
<td>$\Delta \gamma_1 + \Delta \gamma_2$</td>
<td>-0.0041</td>
<td>0.0319</td>
<td>-0.0070**</td>
<td>0.0389</td>
</tr>
<tr>
<td></td>
<td>(0.0114)</td>
<td>(0.0178)</td>
<td>(0.0148)</td>
<td>(0.0231)</td>
</tr>
<tr>
<td>$\Delta \gamma_1 + \Delta \gamma_2$</td>
<td>-0.0208</td>
<td>0.0288</td>
<td>-0.0465</td>
<td>0.0753***</td>
</tr>
<tr>
<td></td>
<td>(0.0128)</td>
<td>(0.0199)</td>
<td>(0.0167)</td>
<td>(0.0260)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of FE estimates of equation (1) in the paper. The outcome where $STEM_{i2}$ denotes a dummy taking the value 1 if student $i$ completes a STEM field University degree and 0 otherwise. $X$ includes the following covariates: region of residence, social class and parents’ jobs. The notation used in the table reflects the formulation presented in equation (5) in the paper. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table 12: Heterogeneity of the effects of parental education on STEM degree choice at the University wrt to the same choice at High School, pooled and by gender of the student. Fixed effects (FE) estimates.

<table>
<thead>
<tr>
<th></th>
<th>Pooled</th>
<th>Males</th>
<th>Females</th>
<th>Males-Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Both parents with</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>degree in STEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta \phi_{11}$</td>
<td>0.0322**</td>
<td>0.1093***</td>
<td>-0.0134</td>
<td>0.1227***</td>
</tr>
<tr>
<td></td>
<td>(0.0154)</td>
<td>(0.0222)</td>
<td>(0.0215)</td>
<td>(0.0309)</td>
</tr>
<tr>
<td>Father with degree in STEM different mother education levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta \phi_{12}$</td>
<td>0.0444***</td>
<td>0.0813***</td>
<td>0.0339**</td>
<td>0.0474**</td>
</tr>
<tr>
<td></td>
<td>(0.0113)</td>
<td>(0.0170)</td>
<td>(0.0153)</td>
<td>(0.0228)</td>
</tr>
<tr>
<td>$\Delta \phi_{13}$</td>
<td>-0.0042</td>
<td>0.0387**</td>
<td>-0.0221</td>
<td>0.0608***</td>
</tr>
<tr>
<td></td>
<td>(0.0104)</td>
<td>(0.0156)</td>
<td>(0.0139)</td>
<td>(0.0209)</td>
</tr>
<tr>
<td>$\Delta \phi_{14}$</td>
<td>-0.0351</td>
<td>0.0014</td>
<td>-0.0445</td>
<td>0.0459</td>
</tr>
<tr>
<td></td>
<td>(0.0265)</td>
<td>(0.0388)</td>
<td>(0.0362)</td>
<td>(0.0531)</td>
</tr>
<tr>
<td>Mother with degree in STEM different father education levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta \phi_{21}$</td>
<td>-0.0826***</td>
<td>-0.0502**</td>
<td>-0.0881***</td>
<td>0.0379</td>
</tr>
<tr>
<td></td>
<td>(0.0151)</td>
<td>(0.0230)</td>
<td>(0.0199)</td>
<td>(0.0304)</td>
</tr>
<tr>
<td>$\Delta \phi_{31}$</td>
<td>-0.0504***</td>
<td>0.0039</td>
<td>-0.0774***</td>
<td>0.0813***</td>
</tr>
<tr>
<td></td>
<td>(0.0152)</td>
<td>(0.0228)</td>
<td>(0.0205)</td>
<td>(0.0306)</td>
</tr>
<tr>
<td>$\Delta \phi_{41}$</td>
<td>-0.0076</td>
<td>0.0331</td>
<td>-0.0278</td>
<td>0.0609</td>
</tr>
<tr>
<td></td>
<td>(0.0336)</td>
<td>(0.0549)</td>
<td>(0.0423)</td>
<td>(0.0694)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of FE estimates of equation (3) in the paper on the pooled sample, on the sample of male students and on the sample of females students, respectively. The outcome is $STEM_i$, a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise. All equations include $X$, the following set of covariates: region of residence, social class and parents’ jobs. The notation used in the table mimics the one used in the formulation of the baseline case presented in equation (5) in the paper. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table 13: Estimates of the effects of parental education on STEM degree choice at the University, pooled and by gender of the student controlling for field choice in High School (HS)

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(mean: 0.231)</td>
<td>(mean: 0.348)</td>
<td>(mean: 0.155)</td>
<td></td>
</tr>
<tr>
<td><strong>STEM choice in HS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_{HS, stem}$</td>
<td>0.2473***</td>
<td>0.2822***</td>
<td>0.1666***</td>
<td>0.1156***</td>
</tr>
<tr>
<td></td>
<td>(0.0020)</td>
<td>(0.0035)</td>
<td>(0.0026)</td>
<td>(0.0044)</td>
</tr>
<tr>
<td><strong>Father education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.1287***</td>
<td>0.1337***</td>
<td>0.1202***</td>
<td>0.0135</td>
</tr>
<tr>
<td></td>
<td>(0.0058)</td>
<td>(0.0095)</td>
<td>(0.0072)</td>
<td>(0.0119)</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>-0.0233***</td>
<td>-0.0539***</td>
<td>-0.0065</td>
<td>-0.0474***</td>
</tr>
<tr>
<td></td>
<td>(0.0042)</td>
<td>(0.0073)</td>
<td>(0.0048)</td>
<td>(0.0087)</td>
</tr>
<tr>
<td>$\delta_3$</td>
<td>0.0136***</td>
<td>0.0137***</td>
<td>0.0099***</td>
<td>0.0038</td>
</tr>
<tr>
<td></td>
<td>(0.0026)</td>
<td>(0.0050)</td>
<td>(0.0028)</td>
<td>(0.0057)</td>
</tr>
<tr>
<td><strong>Mother education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.0710***</td>
<td>0.0704***</td>
<td>0.0720***</td>
<td>-0.0016</td>
</tr>
<tr>
<td></td>
<td>(0.0072)</td>
<td>(0.0117)</td>
<td>(0.0089)</td>
<td>(0.0146)</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>0.0150***</td>
<td>0.0043</td>
<td>0.0169***</td>
<td>-0.0126</td>
</tr>
<tr>
<td></td>
<td>(0.0042)</td>
<td>(0.0074)</td>
<td>(0.0048)</td>
<td>(0.0088)</td>
</tr>
<tr>
<td>$\delta_3$</td>
<td>0.0045*</td>
<td>-0.0074</td>
<td>0.0088***</td>
<td>-0.0162***</td>
</tr>
<tr>
<td></td>
<td>(0.0027)</td>
<td>(0.0051)</td>
<td>(0.0029)</td>
<td>(0.0059)</td>
</tr>
<tr>
<td>$\delta_1 + \delta_2$</td>
<td>0.1054***</td>
<td>0.0798***</td>
<td>0.1136***</td>
<td>-0.0339</td>
</tr>
<tr>
<td></td>
<td>(0.0085)</td>
<td>(0.0146)</td>
<td>(0.0101)</td>
<td>(0.0177)</td>
</tr>
<tr>
<td>$\delta_1 + \delta_2$</td>
<td>0.0860***</td>
<td>0.0747***</td>
<td>0.0889***</td>
<td>-0.0143</td>
</tr>
<tr>
<td></td>
<td>(0.0096)</td>
<td>(0.0163)</td>
<td>(0.0114)</td>
<td>(0.0199)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of OLS estimates of equation (10) in the paper. The outcome variable $STEM_{i2}$ denotes a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise; $FEdu_i$, $MEdu_{j,i,j} \in \{1,2,3,4\}$ are dummy variables denoting the qualification level of fathers and mothers respectively, where $i = 1$ ($j = 1$) if the father (mother) has STEM-degree qualification, $i = 2$ ($j = 2$) if the father (mother) has a non-STEM degree qualification, $i = 3$ ($j = 3$) if the father (mother) has a high school qualification and $i = 4$ ($j = 4$) if the father (mother) has a junior high school (JHS) qualification or less. $X$ includes the following covariates: region of residence, social class and parents’ jobs. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table 14: Estimates of the effects of parental education on STEM degree choice at the University, pooled and by gender of the student controlling for field choice in High School (HS)

<table>
<thead>
<tr>
<th></th>
<th>Pooled (mean: 0.231)</th>
<th>Males (mean: 0.348)</th>
<th>Females (mean: 0.155)</th>
<th>Males-Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>STEM choice in HS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_{HS,stem}$</td>
<td>0.2473***</td>
<td>0.2823***</td>
<td>0.1665***</td>
<td>0.1157***</td>
</tr>
<tr>
<td></td>
<td>(0.0020)</td>
<td>(0.0035)</td>
<td>(0.0026)</td>
<td>(0.0044)</td>
</tr>
<tr>
<td>Both parents with degree in STEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_{11}$</td>
<td>0.1974***</td>
<td>0.2013***</td>
<td>0.1915***</td>
<td>0.0098</td>
</tr>
<tr>
<td></td>
<td>(0.0122)</td>
<td>(0.0187)</td>
<td>(0.0160)</td>
<td>(0.0246)</td>
</tr>
<tr>
<td>Father with degree in STEM, different mother ed. levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_{12}$</td>
<td>0.1500***</td>
<td>0.1439***</td>
<td>0.1425***</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td>(0.0088)</td>
<td>(0.0141)</td>
<td>(0.0111)</td>
<td>(0.0180)</td>
</tr>
<tr>
<td>$\theta_{13}$</td>
<td>0.1353***</td>
<td>0.1319***</td>
<td>0.1320***</td>
<td>-0.0001</td>
</tr>
<tr>
<td></td>
<td>(0.0082)</td>
<td>(0.0133)</td>
<td>(0.0101)</td>
<td>(0.0167)</td>
</tr>
<tr>
<td>$\theta_{14}$</td>
<td>0.1023***</td>
<td>0.1078***</td>
<td>0.0836***</td>
<td>0.0242</td>
</tr>
<tr>
<td></td>
<td>(0.0210)</td>
<td>(0.0335)</td>
<td>(0.0256)</td>
<td>(0.0422)</td>
</tr>
<tr>
<td>Mother with degree in STEM, different father ed. levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_{21}$</td>
<td>0.0231**</td>
<td>-0.0049</td>
<td>0.0376***</td>
<td>-0.0425*</td>
</tr>
<tr>
<td></td>
<td>(0.0109)</td>
<td>(0.0178)</td>
<td>(0.0134)</td>
<td>(0.0223)</td>
</tr>
<tr>
<td>$\theta_{31}$</td>
<td>0.1045***</td>
<td>0.1058***</td>
<td>0.1016***</td>
<td>0.0043</td>
</tr>
<tr>
<td></td>
<td>(0.0119)</td>
<td>(0.0192)</td>
<td>(0.0146)</td>
<td>(0.0241)</td>
</tr>
<tr>
<td>$\theta_{41}$</td>
<td>0.1388***</td>
<td>0.1548***</td>
<td>0.1313***</td>
<td>0.0235</td>
</tr>
<tr>
<td></td>
<td>(0.0266)</td>
<td>(0.0460)</td>
<td>(0.0316)</td>
<td>(0.0558)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of OLS estimates on the pooled sample, on the sample of male students and on the sample of female students, respectively, in the same spirit of equation (3) in the paper. The outcome is $STEM_i$, a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise. All equations include $X$, the following set of covariates: region of residence, social class and parents’ jobs. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. $\theta_{ij}$ denote the differential effect of having a father with qualification $i$ and a mother with qualification $j$ on the probability of completing a STEM University degree, respectively, with respect to the case in which both parents hold a qualification equal or lower to a junior high school degree. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Figure 1: Direct and indirect effects of parental degree on STEM field choice at University, pooled and by gender of the student.

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. The total effect at University presents the estimates of parameters $\alpha_i$ and $\alpha_j$ of equation (1) in the paper. The direct effect at University presents the estimates of parameters $\delta_i$ and $\delta_j$ of equation (10) in the paper. The indirect effect presents the estimates of parameters $\delta_{HS,stem} \lambda_i$ and $\delta_{HS,stem} \lambda_j$ of equations (10) and (2) in the paper.
Figure 2: Direct and indirect effects of parental degree on STEM field choice at University; males students only.

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. The total effect at University presents the estimates of parameters $\alpha_i$ and $\alpha_j$ of equation (1) in the paper. The direct effect at University presents the estimates of parameters $\delta_i$ and $\delta_j$ of equation (10) in the paper. The indirect effect presents the estimates of parameters $\delta_{HS_{stem}} \lambda_i$ and $\delta_{HS_{stem}} \lambda_j$ of equations (10) and (2) in the paper.
Figure 3: Direct and indirect effects of parental degree on STEM field choice at University; female students only.

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. The total effect at University presents the estimates of parameters $\alpha_i$ and $\alpha_j$ of equation (1) in the paper. The direct effect at University presents the estimates of parameters $\delta_i$ and $\delta_j$ of equation (10) in the paper. The indirect effect presents the estimates of parameters $\delta_{HS stemmed_i}$ and $\delta_{HS stemmed_j}$ of equations (10) and (2) in the paper.
Table 15: Effects of parental education on STEM degree choice at the University, pooled and by gender of the student

<table>
<thead>
<tr>
<th>Father education</th>
<th>Pooled</th>
<th>Males</th>
<th>Females</th>
<th>Males-Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.1418***</td>
<td>0.1409***</td>
<td>0.1234***</td>
<td>0.0175</td>
</tr>
<tr>
<td></td>
<td>(0.0070)</td>
<td>(0.0111)</td>
<td>(0.0086)</td>
<td>(0.0140)</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>-0.0045</td>
<td>-0.0356***</td>
<td>0.0053</td>
<td>-0.0409***</td>
</tr>
<tr>
<td></td>
<td>(0.0049)</td>
<td>(0.0085)</td>
<td>(0.0057)</td>
<td>(0.0103)</td>
</tr>
<tr>
<td>( \alpha_3 )</td>
<td>0.0245***</td>
<td>0.0222***</td>
<td>0.0157***</td>
<td>0.0065</td>
</tr>
<tr>
<td></td>
<td>(0.0028)</td>
<td>(0.0054)</td>
<td>(0.0030)</td>
<td>(0.0062)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mother education</th>
<th>Pooled</th>
<th>Males</th>
<th>Females</th>
<th>Males-Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.1037***</td>
<td>0.0919***</td>
<td>0.0977***</td>
<td>-0.0058</td>
</tr>
<tr>
<td></td>
<td>(0.0077)</td>
<td>(0.0124)</td>
<td>(0.0095)</td>
<td>(0.0156)</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>0.0203***</td>
<td>-0.0001</td>
<td>0.0223***</td>
<td>-0.0223**</td>
</tr>
<tr>
<td></td>
<td>(0.0045)</td>
<td>(0.0080)</td>
<td>(0.0051)</td>
<td>(0.0095)</td>
</tr>
<tr>
<td>( \alpha_3 )</td>
<td>0.0140***</td>
<td>-0.0024</td>
<td>0.0151***</td>
<td>-0.0175***</td>
</tr>
<tr>
<td></td>
<td>(0.0028)</td>
<td>(0.0054)</td>
<td>(0.0030)</td>
<td>(0.0062)</td>
</tr>
</tbody>
</table>

Interaction terms:
father education \times liberal professional

| \( \alpha_{L1} \) | 0.0555*** | 0.0386** | 0.0538*** | -0.0151       |
|                  | (0.0122) | (0.0194) | (0.0152) | (0.0246)      |
| \( \alpha_{L2} \) | -0.0387*** | -0.0891*** | -0.0295** | -0.0686***    |
|                  | (0.0076) | (0.0131) | (0.0089) | (0.0158)      |
| \( \alpha_{L3} \) | 0.0044  | -0.0170  | 0.0073  | -0.0243       |
|                  | (0.0077) | (0.0136) | (0.0087) | (0.0161)      |
| \( \alpha_{L4} \) | -0.0083 | -0.0112 | -0.0147 | 0.0035        |
|                  | (0.0091) | (0.0177) | (0.0094) | (0.0200)      |

Interaction terms:
mother education \times liberal professional

| \( \alpha_{L1} \) | -0.0243 | -0.0604* | -0.0001 | -0.0603       |
|                  | (0.0210) | (0.0336) | (0.0261) | (0.0426)      |
| \( \alpha_{L2} \) | -0.0172** | -0.0529*** | -0.0020 | -0.0509***    |
|                  | (0.0084) | (0.0145) | (0.0099) | (0.0175)      |
| \( \alpha_{L3} \) | -0.0185* | -0.0712*** | 0.0104 | -0.0816***    |
|                  | (0.0100) | (0.0177) | (0.0115) | (0.0211)      |
| \( \alpha_{L4} \) | -0.0093 | -0.0666** | 0.0079 | -0.0744**     |
|                  | (0.0175) | (0.0320) | (0.0192) | (0.0373)      |

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. The table reports OLS estimates of linear regressions models that also include following covariates: region of residence, social class and parents’ jobs. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
A Classification of STEM degrees for students

A.1 Detailed classification of STEM fields according to ISCED-F 2013

- **05 Natural sciences, mathematics and statistics**
  - 051 Biological and related sciences
    - 0511 Biology
    - 0512 Biochemistry
  - 052 Environment
    - 0521 Environmental sciences
    - 0522 Natural environments and wildlife
  - 053 Physical sciences
    - 0531 Chemistry
    - 0532 Earth sciences
    - 0533 Physics
  - 054 Mathematics and statistics
    - 0541 Mathematics
    - 0542 Statistics
  - 058 Interdisciplinary programmes involving broad field 05
    - 0588 Interdisciplinary programmes involving broad field 05

- **06 Information and Communication Technologies**
  - 061 Information and Communication Technologies
    - 0611 Computer use
    - 0612 Database and network design and administration
    - 0613 Software and applications development and analysis
  - 068 Interdisciplinary programmes involving broad field 06
    - 0688 Interdisciplinary programmes involving broad field 06

- **07 Engineering, manufacturing and construction**
  - 071 Engineering and engineering trades
0711 Chemical engineering and processes
0712 Environmental protection technology
0713 Electricity and energy
0714 Electronics and automation
0715 Mechanics and metal trades
0716 Motor vehicles, ships and aircraft

- 072 Manufacturing and processing
  0721 Food processing
  0722 Materials (glass, paper, plastic and wood)
  0723 Textiles (clothes, footwear and leather)
  0724 Mining and extraction

- 073 Architecture and construction
  0731 Architecture and town planning
  0732 Building and civil engineering

- 078 Interdisciplinary programmes involving broad field 07
  0788 Interdisciplinary programmes involving broad field 07

## A.2 List of Italian STEM degrees

<table>
<thead>
<tr>
<th>Numero</th>
<th>Denominazione</th>
<th>Ordinamento</th>
<th>ISCED_F.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Biotecnologie</td>
<td>D.M. 509/99</td>
<td>512</td>
</tr>
<tr>
<td>4</td>
<td>Scienze dell’architettura e dell’ingegneria edile</td>
<td>D.M. 509/99</td>
<td>731</td>
</tr>
<tr>
<td>7</td>
<td>Urbanistica e scienze della pianificazione territoriale e ambientale</td>
<td>D.M. 509/99</td>
<td>731</td>
</tr>
<tr>
<td>8</td>
<td>Ingegneria civile e ambientale</td>
<td>D.M. 509/99</td>
<td>732</td>
</tr>
<tr>
<td>9</td>
<td>Ingegneria dell’informazione</td>
<td>D.M. 509/99</td>
<td>688</td>
</tr>
<tr>
<td>10</td>
<td>Ingegneria industriale</td>
<td>D.M. 509/99</td>
<td>719</td>
</tr>
<tr>
<td>12</td>
<td>Scienze biologiche</td>
<td>D.M. 509/99</td>
<td>511</td>
</tr>
<tr>
<td>16</td>
<td>Scienze della Terra</td>
<td>D.M. 509/99</td>
<td>532</td>
</tr>
<tr>
<td>21</td>
<td>Scienze e tecnologie chimiche</td>
<td>D.M. 509/99</td>
<td>531</td>
</tr>
<tr>
<td>22</td>
<td>Scienze e tecnologie della navigazione marittima e aerea</td>
<td>D.M. 509/99</td>
<td>539</td>
</tr>
<tr>
<td>24</td>
<td>Scienze e tecnologie farmaceutiche</td>
<td>D.M. 509/99</td>
<td>531</td>
</tr>
<tr>
<td>25</td>
<td>Scienze e tecnologie fisiche</td>
<td>D.M. 509/99</td>
<td>533</td>
</tr>
<tr>
<td>26</td>
<td>Scienze e tecnologie informatiche</td>
<td>D.M. 509/99</td>
<td>688</td>
</tr>
<tr>
<td>27</td>
<td>Scienze e tecnologie per l’ambiente e la natura</td>
<td>D.M. 509/99</td>
<td>521</td>
</tr>
<tr>
<td>32</td>
<td>Scienze matematiche</td>
<td>D.M. 509/99</td>
<td>541</td>
</tr>
<tr>
<td>37</td>
<td>Scienze statistiche</td>
<td>D.M. 509/99</td>
<td>542</td>
</tr>
<tr>
<td>41</td>
<td>Tecnologie per la conservazione e il restauro dei beni culturali</td>
<td>D.M. 509/99</td>
<td>531</td>
</tr>
<tr>
<td>03/S</td>
<td>Architettura del paesaggio</td>
<td>D.M. 509/99</td>
<td>731</td>
</tr>
<tr>
<td>04/S</td>
<td>Architettura e ingegneria edile</td>
<td>D.M. 509/99</td>
<td>788</td>
</tr>
<tr>
<td>06/S</td>
<td>Biologia</td>
<td>D.M. 509/99</td>
<td>511</td>
</tr>
<tr>
<td>08/S</td>
<td>Biotecnologie industriali</td>
<td>D.M. 509/99</td>
<td>512</td>
</tr>
<tr>
<td>09/S</td>
<td>Biotecnologie mediche, veterinarie e farmaceutiche</td>
<td>D.M. 509/99</td>
<td>512</td>
</tr>
<tr>
<td>10/S</td>
<td>Conservazione dei beni architettonici e ambientali</td>
<td>D.M. 509/99</td>
<td>731</td>
</tr>
<tr>
<td>100/S</td>
<td>Tecniche e metodi per la società dell’informazione</td>
<td>D.M. 509/99</td>
<td>688</td>
</tr>
<tr>
<td>11/S</td>
<td>Conservazione dei beni scientifici e della civilt industriale</td>
<td>D.M. 509/99</td>
<td>588</td>
</tr>
<tr>
<td>Numero</td>
<td>Denominazione</td>
<td>Ordinamento</td>
<td>ISCED_F_3</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>12/S</td>
<td>Conservazione e restauro del patrimonio storico-artistico</td>
<td>D.M. 509/99</td>
<td>588</td>
</tr>
<tr>
<td>20/S</td>
<td>Fisica</td>
<td>D.M. 509/99</td>
<td>533</td>
</tr>
<tr>
<td>23/S</td>
<td>Informatica</td>
<td>D.M. 509/99</td>
<td>688</td>
</tr>
<tr>
<td>25/S</td>
<td>Ingegneria aerospaziale e astronautica</td>
<td>D.M. 509/99</td>
<td>716</td>
</tr>
<tr>
<td>26/S</td>
<td>Ingegneria biomedica</td>
<td>D.M. 509/99</td>
<td>719</td>
</tr>
<tr>
<td>27/S</td>
<td>Ingegneria chimica</td>
<td>D.M. 509/99</td>
<td>711</td>
</tr>
<tr>
<td>28/S</td>
<td>Ingegneria civile</td>
<td>D.M. 509/99</td>
<td>732</td>
</tr>
<tr>
<td>29/S</td>
<td>Ingegneria dell’automazione</td>
<td>D.M. 509/99</td>
<td>714</td>
</tr>
<tr>
<td>30/S</td>
<td>Ingegneria delle telecomunicazioni</td>
<td>D.M. 509/99</td>
<td>714</td>
</tr>
<tr>
<td>31/S</td>
<td>Ingegneria elettrica</td>
<td>D.M. 509/99</td>
<td>713</td>
</tr>
<tr>
<td>32/S</td>
<td>Ingegneria elettronica</td>
<td>D.M. 509/99</td>
<td>714</td>
</tr>
<tr>
<td>33/S</td>
<td>Ingegneria energetica e nucleare</td>
<td>D.M. 509/99</td>
<td>713</td>
</tr>
<tr>
<td>34/S</td>
<td>Ingegneria gestionale</td>
<td>D.M. 509/99</td>
<td>719</td>
</tr>
<tr>
<td>35/S</td>
<td>Ingegneria informatica</td>
<td>D.M. 509/99</td>
<td>688</td>
</tr>
<tr>
<td>36/S</td>
<td>Ingegneria meccanica</td>
<td>D.M. 509/99</td>
<td>715</td>
</tr>
<tr>
<td>37/S</td>
<td>Ingegneria navale</td>
<td>D.M. 509/99</td>
<td>716</td>
</tr>
<tr>
<td>38/S</td>
<td>Ingegneria per l’ambiente e il territorio</td>
<td>D.M. 509/99</td>
<td>712</td>
</tr>
<tr>
<td>45/S</td>
<td>Matematica</td>
<td>D.M. 509/99</td>
<td>541</td>
</tr>
<tr>
<td>48/S</td>
<td>Metodi per l’analisi valutativa dei sistemi complessi</td>
<td>D.M. 509/99</td>
<td>542</td>
</tr>
<tr>
<td>50/S</td>
<td>Modellistica matematico-fisica per l’ingegneria</td>
<td>D.M. 509/99</td>
<td>788</td>
</tr>
<tr>
<td>54/S</td>
<td>Pianificazione territoriale urbanistica e ambientale</td>
<td>D.M. 509/99</td>
<td>731</td>
</tr>
<tr>
<td>61/S</td>
<td>Scienza e ingegneria dei materiali</td>
<td>D.M. 509/99</td>
<td>788</td>
</tr>
<tr>
<td>62/S</td>
<td>Scienze chimiche</td>
<td>D.M. 509/99</td>
<td>531</td>
</tr>
<tr>
<td>66/S</td>
<td>Scienze dell’universo</td>
<td>D.M. 509/99</td>
<td>533</td>
</tr>
<tr>
<td>68/S</td>
<td>Scienze della natura</td>
<td>D.M. 509/99</td>
<td>521</td>
</tr>
<tr>
<td>69/S</td>
<td>Scienze della nutrizione umana</td>
<td>D.M. 509/99</td>
<td>519</td>
</tr>
<tr>
<td>80/S</td>
<td>Scienze e tecnologie dei sistemi di navigazione</td>
<td>D.M. 509/99</td>
<td>539</td>
</tr>
<tr>
<td>81/S</td>
<td>Scienze e tecnologie della chimica industriale</td>
<td>D.M. 509/99</td>
<td>531</td>
</tr>
<tr>
<td>82/S</td>
<td>Scienze e tecnologie per l’ambiente e il territorio</td>
<td>D.M. 509/99</td>
<td>521</td>
</tr>
<tr>
<td>85/S</td>
<td>Scienze geofisiche</td>
<td>D.M. 509/99</td>
<td>532</td>
</tr>
<tr>
<td>86/S</td>
<td>Scienze geologiche</td>
<td>D.M. 509/99</td>
<td>532</td>
</tr>
<tr>
<td>Numero</td>
<td>Denominazione</td>
<td>Ordinamento</td>
<td>ISCED_F_3</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>90/S</td>
<td>Statistica demografica e sociale</td>
<td>D.M. 509/99</td>
<td>542</td>
</tr>
<tr>
<td>91/S</td>
<td>Statistica economica, finanziaria ed attuariale</td>
<td>D.M. 509/99</td>
<td>542</td>
</tr>
<tr>
<td>92/S</td>
<td>Statistica per la ricerca sperimentale</td>
<td>D.M. 509/99</td>
<td>542</td>
</tr>
<tr>
<td>L-02</td>
<td>Biotecnologie</td>
<td>D.M. 270/04</td>
<td>512</td>
</tr>
<tr>
<td>L-07</td>
<td>Ingegneria civile e ambientale</td>
<td>D.M. 270/04</td>
<td>732</td>
</tr>
<tr>
<td>L-08</td>
<td>Ingegneria dell’informazione</td>
<td>D.M. 270/04</td>
<td>688</td>
</tr>
<tr>
<td>L-09</td>
<td>Ingegneria industriale</td>
<td>D.M. 270/04</td>
<td>719</td>
</tr>
<tr>
<td>L-13</td>
<td>Scienze biologiche</td>
<td>D.M. 270/04</td>
<td>511</td>
</tr>
<tr>
<td>L-17</td>
<td>Scienze dell’architettura</td>
<td>D.M. 270/04</td>
<td>731</td>
</tr>
<tr>
<td>L-21</td>
<td>Scienze della pianificazione territoriale, urbanistica, paesaggistica e ambientale</td>
<td>D.M. 270/04</td>
<td>731</td>
</tr>
<tr>
<td>L-23</td>
<td>Scienze e tecniche dell’edilizia</td>
<td>D.M. 270/04</td>
<td>732</td>
</tr>
<tr>
<td>L-27</td>
<td>Scienze e tecnologie chimiche</td>
<td>D.M. 270/04</td>
<td>531</td>
</tr>
<tr>
<td>L-28</td>
<td>Scienze e tecnologie della navigazione</td>
<td>D.M. 270/04</td>
<td>539</td>
</tr>
<tr>
<td>L-29</td>
<td>Scienze e tecnologie farmaceutiche</td>
<td>D.M. 270/04</td>
<td>531</td>
</tr>
<tr>
<td>L-30</td>
<td>Scienze e tecnologie fisiche</td>
<td>D.M. 270/04</td>
<td>533</td>
</tr>
<tr>
<td>L-31</td>
<td>Scienze e tecnologie informatiche</td>
<td>D.M. 270/04</td>
<td>688</td>
</tr>
<tr>
<td>L-32</td>
<td>Scienze e tecnologie per l’ambiente e la natura</td>
<td>D.M. 270/04</td>
<td>521</td>
</tr>
<tr>
<td>L-34</td>
<td>Scienze geologiche</td>
<td>D.M. 270/04</td>
<td>532</td>
</tr>
<tr>
<td>L-35</td>
<td>Scienze matematiche</td>
<td>D.M. 270/04</td>
<td>541</td>
</tr>
<tr>
<td>L-41</td>
<td>Statistica</td>
<td>D.M. 270/04</td>
<td>542</td>
</tr>
<tr>
<td>L-43</td>
<td>Diagnostica per la conservazione dei beni culturali</td>
<td>D.M. 270/04</td>
<td>588</td>
</tr>
<tr>
<td>LM-03</td>
<td>Architettura del paesaggio</td>
<td>D.M. 270/04</td>
<td>731</td>
</tr>
<tr>
<td>LM-04</td>
<td>Architettura e ingegneria edile-architettura</td>
<td>D.M. 270/04</td>
<td>788</td>
</tr>
<tr>
<td>LM-04cu</td>
<td>Architettura e ingegneria edile-architettura (quinquennale)</td>
<td>D.M. 270/04</td>
<td>788</td>
</tr>
<tr>
<td>LM-06</td>
<td>Biologia</td>
<td>D.M. 270/04</td>
<td>511</td>
</tr>
<tr>
<td>LM-08</td>
<td>Biotecnologie industriali</td>
<td>D.M. 270/04</td>
<td>512</td>
</tr>
<tr>
<td>LM-09</td>
<td>Biotecnologie mediche, veterinarie e farmaceutiche</td>
<td>D.M. 270/04</td>
<td>512</td>
</tr>
<tr>
<td>LM-10</td>
<td>Conservazione dei beni architettonici e ambientali</td>
<td>D.M. 270/04</td>
<td>731</td>
</tr>
<tr>
<td>LM-11</td>
<td>Scienze per la conservazione dei beni culturali</td>
<td>D.M. 270/04</td>
<td>588</td>
</tr>
<tr>
<td>Numero</td>
<td>Denominazione</td>
<td>Ordinamento</td>
<td>ISCED_F_3</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>LM-17</td>
<td>Fisica</td>
<td>D.M. 270/04</td>
<td>533</td>
</tr>
<tr>
<td>LM-18</td>
<td>Informatica</td>
<td>D.M. 270/04</td>
<td>688</td>
</tr>
<tr>
<td>LM-20</td>
<td>Ingegneria aerospaziale e astronautica</td>
<td>D.M. 270/04</td>
<td>716</td>
</tr>
<tr>
<td>LM-21</td>
<td>Ingegneria biomedica</td>
<td>D.M. 270/04</td>
<td>719</td>
</tr>
<tr>
<td>LM-22</td>
<td>Ingegneria chimica</td>
<td>D.M. 270/04</td>
<td>711</td>
</tr>
<tr>
<td>LM-23</td>
<td>Ingegneria civile</td>
<td>D.M. 270/04</td>
<td>732</td>
</tr>
<tr>
<td>LM-24</td>
<td>Ingegneria dei sistemi edilizi</td>
<td>D.M. 270/04</td>
<td>732</td>
</tr>
<tr>
<td>LM-25</td>
<td>Ingegneria dell’automazione</td>
<td>D.M. 270/04</td>
<td>714</td>
</tr>
<tr>
<td>LM-26</td>
<td>Ingegneria della sicurezza</td>
<td>D.M. 270/04</td>
<td>719</td>
</tr>
<tr>
<td>LM-27</td>
<td>Ingegneria delle telecomunicazioni</td>
<td>D.M. 270/04</td>
<td>714</td>
</tr>
<tr>
<td>LM-28</td>
<td>Ingegneria elettrica</td>
<td>D.M. 270/04</td>
<td>713</td>
</tr>
<tr>
<td>LM-29</td>
<td>Ingegneria elettronica</td>
<td>D.M. 270/04</td>
<td>714</td>
</tr>
<tr>
<td>LM-30</td>
<td>Ingegneria energetica e nucleare</td>
<td>D.M. 270/04</td>
<td>713</td>
</tr>
<tr>
<td>LM-31</td>
<td>Ingegneria gestionale</td>
<td>D.M. 270/04</td>
<td>719</td>
</tr>
<tr>
<td>LM-32</td>
<td>Ingegneria informatica</td>
<td>D.M. 270/04</td>
<td>688</td>
</tr>
<tr>
<td>LM-33</td>
<td>Ingegneria meccanica</td>
<td>D.M. 270/04</td>
<td>715</td>
</tr>
<tr>
<td>LM-34</td>
<td>Ingegneria navale</td>
<td>D.M. 270/04</td>
<td>716</td>
</tr>
<tr>
<td>LM-35</td>
<td>Ingegneria per l’ambiente e il territorio</td>
<td>D.M. 270/04</td>
<td>712</td>
</tr>
<tr>
<td>LM-40</td>
<td>Matematica</td>
<td>D.M. 270/04</td>
<td>541</td>
</tr>
<tr>
<td>LM-44</td>
<td>Modellistica matematico-fisica per l’ingegneria</td>
<td>D.M. 270/04</td>
<td>788</td>
</tr>
<tr>
<td>LM-48</td>
<td>Pianificazione territoriale urbanistica e ambientale</td>
<td>D.M. 270/04</td>
<td>731</td>
</tr>
<tr>
<td>LM-53</td>
<td>Scienza e ingegneria dei materiali</td>
<td>D.M. 270/04</td>
<td>788</td>
</tr>
<tr>
<td>LM-54</td>
<td>Scienze chimiche</td>
<td>D.M. 270/04</td>
<td>531</td>
</tr>
<tr>
<td>LM-58</td>
<td>Scienze dell’universo</td>
<td>D.M. 270/04</td>
<td>533</td>
</tr>
<tr>
<td>LM-60</td>
<td>Scienze della natura</td>
<td>D.M. 270/04</td>
<td>521</td>
</tr>
<tr>
<td>LM-61</td>
<td>Scienze della nutrizione umana</td>
<td>D.M. 270/04</td>
<td>519</td>
</tr>
<tr>
<td>LM-66</td>
<td>Sicurezza informatica</td>
<td>D.M. 270/04</td>
<td>612</td>
</tr>
<tr>
<td>LM-71</td>
<td>Scienze e tecnologie della chimica industriale</td>
<td>D.M. 270/04</td>
<td>531</td>
</tr>
<tr>
<td>LM-72</td>
<td>Scienze e tecnologie della navigazione</td>
<td>D.M. 270/04</td>
<td>539</td>
</tr>
<tr>
<td>LM-74</td>
<td>Scienze e tecnologie geologiche</td>
<td>D.M. 270/04</td>
<td>532</td>
</tr>
<tr>
<td>LM-75</td>
<td>Scienze e tecnologie per l’ambiente e il territorio</td>
<td>D.M. 270/04</td>
<td>521</td>
</tr>
<tr>
<td>LM-79</td>
<td>Scienze geofisiche</td>
<td>D.M. 270/04</td>
<td>532</td>
</tr>
<tr>
<td>LM-82</td>
<td>Scienze statistiche</td>
<td>D.M. 270/04</td>
<td>542</td>
</tr>
</tbody>
</table>
### B Classification of STEM degrees for parents

Given that the parents obtained their degrees before the Bologna process and hold the so-called "vecchio ordinamento" degrees, we cannot classify the field of their degrees according to the new "classes" introduced in 1999. To get around this problem, we use the grouping of "classes" into broader areas of studies defined by Almalaurea. The idea behind this strategy is to assign parents' degree name into an area of study instead of a specific class, in order to make parents' and children's education comparable and classifiable according to the same criteria. Overall, there are 16 broad fields of study: scientific, chemical-pharmaceutical, geo-biological, medical, engineering, architecture, agrarian-veterinarian, economical-statistical, political-sociological, legal, literary, linguistics, teaching, psychological, physical education, defense and security. In total, 52,618 titles for father's degree and 48,545 titles for mother's degree were allocated in one of the sixteen above indicated categories.

It is important to keep in mind that this classification follows a different scheme than the ISCED-F 2013, thus the fields of study do not form direct counterparts of the ISCED-F 2013 fields. While some of them can be completely absorbed into ISCED groups and directly classifiable as STEM or non-STEM, the majority of them cannot. For instance, Engineering and Architecture groups can be entirely classified as STEM since the courses belonging to these groups correspond to ISCED-F 07 Engineering, manufacturing and construction category. The scientific group is also entirely classifiable as STEM although it contains some courses that belong to ISCED-F 05 Natural sciences, mathematics and statistics group, and some courses that form the 06 Information and Communication Technologies group. The biological field is entirely classified as STEM with the exception of one class (Agricultural biotechnology). The chemical courses from the chemical-pharmaceutical group are considered STEM since are inserted in ISCED-F 05 group while the pharmaceutical courses are part of the Medicine group in the ISCED-F classification, thus considered non-STEM by the European Commission. All the remaining areas of study are instead overall classifiable as non-STEM with only few exceptions that fall into STEM groups: the Statistics courses from the economical group and the course in Nutritional sciences from the medical group. In order to have a homogeneous picture of STEM and non-STEM groups, we decided to allocate this isolated classes into more solid STEM groups as follows: Chemical and Statistics classes were assigned to the Scientific group, Nutritional sciences was assigned to the Geo-biological group. As a result, we ended up with four STEM groups that are equivalent to the three ISCED-F STEM groups as illustrated below:

- **Geo-biological**
  - (05) Natural Sciences, Mathematics and Statistics
- **Scientific**
  - (06) Information and Communication Technology
- **Engineering**
  - (07) Engineering, Manufacturing and Construction
- **Architecture**

---

2Before the Bologna process all the study cycle lasted 4 years while after it was separated into 3+2 of Bachelor plus Master. We do not know the date in which parents graduate but given the years in which the Bologna process started (1999) is it highly implausible that parents graduated under the new system.
## C Online Appendix: Testing for heterogeneous effects

Table Online Appendix - 1: Heterogeneity of the effects of parental education on STEM degree choice at the University, pooled and by gender of the student and of the parents.

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1) mean: 0.231</th>
<th>Males (2) mean: 0.348</th>
<th>Females (3) mean: 0.155</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterogeneity by field of parental University degree and gender of the parent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1 - \alpha_2$</td>
<td>0.1783*** (0.0057)</td>
<td>0.2195*** (0.0088)</td>
<td>0.1436*** (0.0071)</td>
<td>0.0759*** (0.0113)</td>
</tr>
<tr>
<td>$\alpha_1 - \alpha_2$</td>
<td>0.0819*** (0.0070)</td>
<td>0.0902** (0.0109)</td>
<td>0.0751*** (0.0087)</td>
<td>0.0151</td>
</tr>
<tr>
<td>$(\alpha_1 - \alpha_2) - (\alpha_1 - \alpha_2)$</td>
<td>0.0964*** (0.0095)</td>
<td>0.1293*** (0.0149)</td>
<td>0.0685*** (0.0119)</td>
<td>0.0608*** (0.0190)</td>
</tr>
<tr>
<td>Heterogeneity by gender of the parent for given level of parental education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1 - \alpha_1$</td>
<td>0.0583*** (0.0105)</td>
<td>0.0676*** (0.0171)</td>
<td>0.0451*** (0.0128)</td>
<td>0.0225</td>
</tr>
<tr>
<td>$\alpha_2 - \alpha_2$</td>
<td>-0.0381*** (0.0070)</td>
<td>-0.0617*** (0.0123)</td>
<td>-0.0234** (0.0079)</td>
<td>-0.0383** (0.0147)</td>
</tr>
<tr>
<td>$\alpha_3 - \alpha_3$</td>
<td>0.0116** (0.0044)</td>
<td>0.0243*** (0.0084)</td>
<td>0.0020</td>
<td>0.0222* (0.0096)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present differences between sets of OLS estimates of equation (1) in the paper, reported in Table 7. The outcome $STEM_2$ denotes a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise; $FEdu_{i, j}$, $MEdu_{i, j}$, $i, j \in \{1, 2, 3, 4\}$ are dummy variables denoting the qualification level of fathers and mothers respectively, where $i = 1$ ($j = 1$) if the father (mother) has STEM-degree qualification, $i = 2$ ($j = 2$) if the father (mother) has a non-STEM degree qualification, $i = 3$ ($j = 3$) if the father (mother) has a high school qualification and $i = 4$ ($j = 4$) if the father (mother) has a junior high school (JHS) qualification or less. $X$ includes the following covariates: region of residence, social class and parents’ jobs. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to differences of interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table Online Appendix - 2: Heterogeneity of the effects of parental education on STEM degree choice at the University, pooled and by gender of the student and of the parents. Interdependent effects

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{11} - \beta_{12}$</td>
<td>0.0670***</td>
<td>0.0690***</td>
<td>0.0683***</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td>(0.0144)</td>
<td>(0.0214)</td>
<td>(0.0188)</td>
<td>(0.0285)</td>
</tr>
<tr>
<td>$\beta_{11} - \beta_{13}$</td>
<td>0.0705***</td>
<td>0.0689***</td>
<td>0.0696***</td>
<td>-0.0007</td>
</tr>
<tr>
<td></td>
<td>(0.0141)</td>
<td>(0.0212)</td>
<td>(0.0184)</td>
<td>(0.0281)</td>
</tr>
<tr>
<td>$\beta_{11} - \beta_{14}$</td>
<td>0.1042***</td>
<td>0.0879*</td>
<td>0.1232***</td>
<td>-0.0354</td>
</tr>
<tr>
<td></td>
<td>(0.0247)</td>
<td>(0.0381)</td>
<td>(0.0304)</td>
<td>(0.0488)</td>
</tr>
<tr>
<td>$\beta_{21} - \beta_{22}$</td>
<td>0.0519***</td>
<td>0.0601***</td>
<td>0.0429**</td>
<td>0.0172</td>
</tr>
<tr>
<td></td>
<td>(0.0111)</td>
<td>(0.0174)</td>
<td>(0.0138)</td>
<td>(0.0222)</td>
</tr>
<tr>
<td>$\beta_{21} - \beta_{23}$</td>
<td>0.0570***</td>
<td>0.0705***</td>
<td>0.0453***</td>
<td>0.0251</td>
</tr>
<tr>
<td></td>
<td>(0.0112)</td>
<td>(0.0177)</td>
<td>(0.0139)</td>
<td>(0.0225)</td>
</tr>
<tr>
<td>$\beta_{21} - \beta_{24}$</td>
<td>0.0757***</td>
<td>0.0991***</td>
<td>0.0517**</td>
<td>0.0474</td>
</tr>
<tr>
<td></td>
<td>(0.0149)</td>
<td>(0.0248)</td>
<td>(0.0180)</td>
<td>(0.0306)</td>
</tr>
<tr>
<td>$\beta_{31} - \beta_{32}$</td>
<td>0.1121***</td>
<td>0.1306***</td>
<td>0.0966***</td>
<td>0.0341</td>
</tr>
<tr>
<td></td>
<td>(0.0124)</td>
<td>(0.0198)</td>
<td>(0.0152)</td>
<td>(0.0250)</td>
</tr>
<tr>
<td>$\beta_{31} - \beta_{33}$</td>
<td>0.1157***</td>
<td>0.1065***</td>
<td>0.1149***</td>
<td>0.0281</td>
</tr>
<tr>
<td></td>
<td>(0.0119)</td>
<td>(0.0188)</td>
<td>(0.0150)</td>
<td>(0.0238)</td>
</tr>
<tr>
<td>$\beta_{31} - \beta_{34}$</td>
<td>0.1218***</td>
<td>0.1239***</td>
<td>0.1058***</td>
<td>-0.0084</td>
</tr>
<tr>
<td></td>
<td>(0.0123)</td>
<td>(0.0199)</td>
<td>(0.0146)</td>
<td>(0.0249)</td>
</tr>
<tr>
<td>$\beta_{41} - \beta_{42}$</td>
<td>0.1617***</td>
<td>0.1822***</td>
<td>0.1511***</td>
<td>0.0311</td>
</tr>
<tr>
<td></td>
<td>(0.0289)</td>
<td>(0.0485)</td>
<td>(0.0340)</td>
<td>(0.0592)</td>
</tr>
<tr>
<td>$\beta_{41} - \beta_{43}$</td>
<td>0.1683***</td>
<td>0.1961*</td>
<td>0.1436***</td>
<td>0.0525</td>
</tr>
<tr>
<td></td>
<td>(0.0275)</td>
<td>(0.0456)</td>
<td>(0.0325)</td>
<td>(0.0560)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present differences between sets of OLS estimates of equation (3) in the paper, reported in Table 2. The outcome is $STEM_{it}$ a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise. All equations include $X$, the following set of covariates: region of residence, social class and parents’ jobs. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to differences between interaction terms of a fully interacted model specification by student gender. $\beta_{ij} i,j \in \{1, 2, 3, 4\}$ denote the differential effect of having a father with qualification $i$ and a mother with qualification $j$ on the probability of completing a STEM University degree, respectively, with respect to the case in which both parents hold a qualification equal or lower to a junior high school degree. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table Online Appendix - 3: Heterogeneity of the effects of parental education on STEM degree choice at High School (HS), pooled and by gender of the student and of the parents.

<table>
<thead>
<tr>
<th></th>
<th>Pooled Males</th>
<th>Females</th>
<th>Males-Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) (mean: 0.518)</td>
<td>(2) (mean: 0.700)</td>
<td>(3) (mean: 0.400)</td>
</tr>
<tr>
<td><strong>Heterogeneity by field of parental University degree and gender of the parent</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_1 - \lambda_2$</td>
<td>0.1067*** (0.0060)</td>
<td>0.1130*** (0.0080)</td>
<td>0.1015*** (0.0083)</td>
</tr>
<tr>
<td>$\lambda_1 - \lambda_2$</td>
<td>0.1046*** (0.0073)</td>
<td>0.0853*** (0.0096)</td>
<td>0.1197 (0.0102)</td>
</tr>
<tr>
<td>$(\lambda_1 - \lambda_2) - (\lambda_1 - \lambda_2)$</td>
<td>0.0021 (0.0100)</td>
<td>0.0276* (0.0132)</td>
<td>-0.0183 (0.0139)</td>
</tr>
<tr>
<td><strong>Heterogeneity by gender of the parent for given level of parental education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_1 - \lambda_1$</td>
<td>0.0028 (0.0112)</td>
<td>0.0152 (0.0151)</td>
<td>-0.0180*** (0.0152)</td>
</tr>
<tr>
<td>$\lambda_2 - \lambda_2$</td>
<td>0.0007 (0.0084)</td>
<td>-0.0124 (0.0120)</td>
<td>0.0003 (0.0108)</td>
</tr>
<tr>
<td>$\lambda_3 - \lambda_3$</td>
<td>0.0098 (0.0053)</td>
<td>0.0111 (0.0080)</td>
<td>0.0057*** (0.0065)</td>
</tr>
</tbody>
</table>

**Notes:** Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present differences between sets of OLS estimates of equation (2) in the paper, reported in Table 9. The outcome $STEM_i$ denotes a dummy taking the value 1 if student $i$ completes a STEM field high school degree and 0 otherwise; $FEdu_i$, $MEdu_{ij}, i,j \in \{1,2,3,4\}$ are dummy variables denoting the qualification level of fathers and mothers respectively, where $i = 1$ ($j = 1$) if the father (mother) has STEM-degree qualification, $i = 2$ ($j = 2$) if the father (mother) has a non-STEM degree qualification, $i = 3$ ($j = 3$) if the father (mother) has a high school qualification and $i = 4$ ($j = 4$) if the father (mother) has a junior high school (JHS) qualification or less. $X$ includes the following covariates: region of residence, social class and parents’ jobs. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to differences between interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table Online Appendix - 4: Heterogeneity of the effects of parental education on STEM degree choice at High School, pooled and by gender of the student and of the parents. Interdependent effects.

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_{11} - \psi_{12}$</td>
<td>0.0792***</td>
<td>0.0410*</td>
<td>0.1155***</td>
<td>-0.0746**</td>
</tr>
<tr>
<td></td>
<td>(0.0138)</td>
<td>(0.0179)</td>
<td>(0.0198)</td>
<td>(0.0267)</td>
</tr>
<tr>
<td>$\psi_{11} - \psi_{13}$</td>
<td>0.0340**</td>
<td>-0.0016</td>
<td>0.0609**</td>
<td>-0.0625*</td>
</tr>
<tr>
<td></td>
<td>(0.0135)</td>
<td>(0.0174)</td>
<td>(0.0194)</td>
<td>(0.0260)</td>
</tr>
<tr>
<td>$\psi_{11} - \psi_{14}$</td>
<td>0.0368</td>
<td>-0.0200</td>
<td>0.0921***</td>
<td>-0.1121*</td>
</tr>
<tr>
<td></td>
<td>(0.0247)</td>
<td>(0.0302)</td>
<td>(0.0349)</td>
<td>(0.0461)</td>
</tr>
<tr>
<td>$\psi_{21} - \psi_{22}$</td>
<td>0.1109***</td>
<td>0.1037***</td>
<td>0.1132***</td>
<td>-0.0094</td>
</tr>
<tr>
<td></td>
<td>(0.0128)</td>
<td>(0.0173)</td>
<td>(0.0179)</td>
<td>(0.0249)</td>
</tr>
<tr>
<td>$\psi_{21} - \psi_{23}$</td>
<td>0.0644***</td>
<td>0.0484**</td>
<td>0.0750***</td>
<td>-0.0267</td>
</tr>
<tr>
<td></td>
<td>(0.0130)</td>
<td>(0.0175)</td>
<td>(0.0181)</td>
<td>(0.0252)</td>
</tr>
<tr>
<td>$\psi_{21} - \psi_{24}$</td>
<td>0.1071***</td>
<td>0.0522*</td>
<td>0.1299***</td>
<td>-0.0777*</td>
</tr>
<tr>
<td></td>
<td>(0.0183)</td>
<td>(0.0260)</td>
<td>(0.0242)</td>
<td>(0.0355)</td>
</tr>
<tr>
<td>$\psi_{31} - \psi_{32}$</td>
<td>0.1213***</td>
<td>0.1047***</td>
<td>0.1322***</td>
<td>-0.0275</td>
</tr>
<tr>
<td></td>
<td>(0.0126)</td>
<td>(0.0163)</td>
<td>(0.0176)</td>
<td>(0.0240)</td>
</tr>
<tr>
<td>$\psi_{31} - \psi_{33}$</td>
<td>0.1219***</td>
<td>0.0990***</td>
<td>0.1370***</td>
<td>-0.0430</td>
</tr>
<tr>
<td></td>
<td>(0.0118)</td>
<td>(0.0163)</td>
<td>(0.0166)</td>
<td>(0.0225)</td>
</tr>
<tr>
<td>$\psi_{31} - \psi_{34}$</td>
<td>0.1500***</td>
<td>0.0940*</td>
<td>0.1631***</td>
<td>-0.0641**</td>
</tr>
<tr>
<td></td>
<td>(0.0125)</td>
<td>(0.0151)</td>
<td>(0.0173)</td>
<td>(0.0237)</td>
</tr>
<tr>
<td>$\psi_{41} - \psi_{42}$</td>
<td>0.1501***</td>
<td>0.1407***</td>
<td>0.1603***</td>
<td>-0.0196</td>
</tr>
<tr>
<td></td>
<td>(0.0293)</td>
<td>(0.0374)</td>
<td>(0.0387)</td>
<td>(0.0538)</td>
</tr>
<tr>
<td>$\psi_{41} - \psi_{43}$</td>
<td>0.1509***</td>
<td>0.1407***</td>
<td>0.1474***</td>
<td>-0.0067</td>
</tr>
<tr>
<td></td>
<td>(0.0271)</td>
<td>(0.0338)</td>
<td>(0.0362)</td>
<td>(0.0495)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present differences between sets of OLS estimates of equation (4) in the paper, reported in Table 10. The outcome is STEM$_i$, a dummy taking the value 1 if student $i$ graduates from a STEM field at high school and 0 otherwise. All equations include $X$, the following set of covariates: region of residence, social class and parents’ jobs. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to differences between interaction terms of a fully interacted model specification by student gender. $\psi_{ij}$, $i,j \in \{1,2,3,4\}$ denote the differential effect of having a father with qualification $i$ and a mother with qualification $j$ on the probability of completing a STEM high-school degree, respectively, with respect to the case in which both parents hold a qualification equal or lower to a junior high school degree. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table Online Appendix - 5: Heterogeneity of the differential effects of parental education on STEM degree choice at the University with respect to STEM field choice at high school (HS), pooled and by gender of the student and of the parents.

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(mean: 0.2314)</td>
<td>(mean: 0.3482)</td>
<td>(mean: 0.1552)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity by field of parental University degree and gender of the parent

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Male-Female</th>
<th>Male-Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \gamma_1 - \Delta \gamma_2$</td>
<td>0.0716***</td>
<td>0.1065***</td>
<td>0.0422***</td>
<td>0.0644***</td>
</tr>
<tr>
<td></td>
<td>(0.0072)</td>
<td>(0.0105)</td>
<td>(0.0098)</td>
<td>(0.0144)</td>
</tr>
<tr>
<td>$\Delta \gamma_1 - \Delta \gamma_2$</td>
<td>-0.0227**</td>
<td>0.0048</td>
<td>-0.0447***</td>
<td>0.0495**</td>
</tr>
<tr>
<td></td>
<td>(0.0089)</td>
<td>(0.0132)</td>
<td>(0.0121)</td>
<td>(0.0179)</td>
</tr>
<tr>
<td>$(\Delta \gamma_1 - \Delta \gamma_2) - (\Delta \gamma_1 - \Delta \gamma_2)$</td>
<td>0.0943***</td>
<td>0.1017***</td>
<td>0.0868***</td>
<td>0.0148</td>
</tr>
<tr>
<td></td>
<td>(0.0121)</td>
<td>(0.0179)</td>
<td>(0.0163)</td>
<td>(0.0242)</td>
</tr>
</tbody>
</table>

Heterogeneity by gender of the parent for given level of parental education

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Male-Female</th>
<th>Male-Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \gamma_1 - \Delta \gamma_1$</td>
<td>0.0555***</td>
<td>0.0524**</td>
<td>0.0632***</td>
<td>-0.0108</td>
</tr>
<tr>
<td></td>
<td>(0.0133)</td>
<td>(0.0202)</td>
<td>(0.0177)</td>
<td>(0.0268)</td>
</tr>
<tr>
<td>$\Delta \gamma_2 - \Delta \gamma_2$</td>
<td>-0.0388***</td>
<td>-0.0493***</td>
<td>-0.0237*</td>
<td>-0.0256</td>
</tr>
<tr>
<td></td>
<td>(0.0093)</td>
<td>(0.0147)</td>
<td>(0.0120)</td>
<td>(0.0190)</td>
</tr>
<tr>
<td>$\Delta \gamma_3 - \Delta \gamma_3$</td>
<td>0.0017</td>
<td>0.0131</td>
<td>-0.0037</td>
<td>0.0168</td>
</tr>
<tr>
<td></td>
<td>(0.0057)</td>
<td>(0.0097)</td>
<td>(0.0071)</td>
<td>(0.0120)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present differences between sets of FE estimates of equation (1) in the paper, reported in Table 11. The outcome where $STEM_i$ denotes a dummy taking the value 1 if student $i$ completes a STEM field University degree and 0 otherwise. $X$ includes the following covariates: region of residence, social class and parents’ jobs. The notation used in the table reflects the formulation presented in equation (5) in the paper. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to differences between interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table Online Appendix - 6: Heterogeneity of the differential effects of parental education on STEM degree choice at the University with respect to STEM field choice at high school (HS), pooled and by gender of the student and of the parents. Interdependent effects.

<table>
<thead>
<tr>
<th></th>
<th>Pooled</th>
<th>Males</th>
<th>Females</th>
<th>Males-Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>$\Delta \phi_{11} - \Delta \phi_{12}$</td>
<td>-0.0122</td>
<td>0.0280</td>
<td>-0.0473</td>
<td>0.0753*</td>
</tr>
<tr>
<td></td>
<td>(0.0176)</td>
<td>(0.0250)</td>
<td>(0.0246)</td>
<td>(0.0351)</td>
</tr>
<tr>
<td>$\Delta \phi_{11} - \Delta \phi_{13}$</td>
<td>0.0365*</td>
<td>0.0706**</td>
<td>0.0087</td>
<td>0.0618</td>
</tr>
<tr>
<td></td>
<td>(0.0172)</td>
<td>(0.0245)</td>
<td>(0.0242)</td>
<td>(0.0344)</td>
</tr>
<tr>
<td>$\Delta \phi_{11} - \Delta \phi_{14}$</td>
<td>0.0673*</td>
<td>0.1079**</td>
<td>0.0311</td>
<td>0.0767</td>
</tr>
<tr>
<td></td>
<td>(0.0300)</td>
<td>(0.0434)</td>
<td>(0.0414)</td>
<td>(0.0600)</td>
</tr>
<tr>
<td>$\Delta \phi_{21} - \Delta \phi_{22}$</td>
<td>-0.0590***</td>
<td>-0.0437</td>
<td>-0.0703***</td>
<td>0.0266</td>
</tr>
<tr>
<td></td>
<td>(0.0152)</td>
<td>(0.0228)</td>
<td>(0.0202)</td>
<td>(0.0305)</td>
</tr>
<tr>
<td>$\Delta \phi_{21} - \Delta \phi_{23}$</td>
<td>-0.0074</td>
<td>0.0221</td>
<td>-0.0297</td>
<td>0.0518</td>
</tr>
<tr>
<td></td>
<td>(0.0154)</td>
<td>(0.0230)</td>
<td>(0.0205)</td>
<td>(0.0308)</td>
</tr>
<tr>
<td>$\Delta \phi_{21} - \Delta \phi_{24}$</td>
<td>-0.0314</td>
<td>0.0470</td>
<td>-0.0782**</td>
<td>0.1251**</td>
</tr>
<tr>
<td></td>
<td>(0.0211)</td>
<td>(0.0323)</td>
<td>(0.0274)</td>
<td>(0.0423)</td>
</tr>
<tr>
<td>$\Delta \phi_{31} - \Delta \phi_{32}$</td>
<td>-0.0093</td>
<td>0.0259</td>
<td>-0.0356</td>
<td>0.0615*</td>
</tr>
<tr>
<td></td>
<td>(0.0157)</td>
<td>(0.0232)</td>
<td>(0.0211)</td>
<td>(0.0314)</td>
</tr>
<tr>
<td>$\Delta \phi_{31} - \Delta \phi_{33}$</td>
<td>-0.0062</td>
<td>0.0299</td>
<td>-0.0482*</td>
<td>0.0611</td>
</tr>
<tr>
<td></td>
<td>(0.0148)</td>
<td>(0.0219)</td>
<td>(0.0208)</td>
<td>(0.0297)</td>
</tr>
<tr>
<td>$\Delta \phi_{31} - \Delta \phi_{34}$</td>
<td>-0.0282</td>
<td>0.0075</td>
<td>-0.0312</td>
<td>0.0557</td>
</tr>
<tr>
<td></td>
<td>(0.0154)</td>
<td>(0.0231)</td>
<td>(0.0201)</td>
<td>(0.0311)</td>
</tr>
<tr>
<td>$\Delta \phi_{41} - \Delta \phi_{42}$</td>
<td>0.0116</td>
<td>0.0415</td>
<td>-0.0092</td>
<td>0.0507</td>
</tr>
<tr>
<td></td>
<td>(0.0356)</td>
<td>(0.0580)</td>
<td>(0.0451)</td>
<td>(0.0735)</td>
</tr>
<tr>
<td>$\Delta \phi_{41} - \Delta \phi_{43}$</td>
<td>0.0174</td>
<td>0.0555</td>
<td>-0.0038</td>
<td>0.0592</td>
</tr>
<tr>
<td></td>
<td>(0.0336)</td>
<td>(0.0549)</td>
<td>(0.0423)</td>
<td>(0.0693)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present differences between sets of FE estimates of equation (3) in the paper, reported in Table 12. The outcome is $STEM_{H2}$ a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise. All equations include $X$, the following set of covariates: region of residence, social class and parents’ jobs. The notation used in the table mimics the one used in the formulation of the baseline case presented in equation (5) in the paper. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to differences between interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table Online Appendix - 7: Heterogeneity of the effects of parental education on STEM degree choice at the University, controlling for STEM high school, pooled and by gender of the student and of the parents.

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Heterogeneity by field of parental University degree and gender of the parent</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_1 - \delta_2$</td>
<td>0.1519***</td>
<td>0.1876***</td>
<td>0.1267***</td>
<td>0.0609***</td>
</tr>
<tr>
<td></td>
<td>(0.0055)</td>
<td>(0.0086)</td>
<td>(0.0070)</td>
<td>(0.0111)</td>
</tr>
<tr>
<td>$\delta_1 - \delta_2$</td>
<td>0.0561***</td>
<td>0.0661***</td>
<td>0.0551***</td>
<td>0.0110</td>
</tr>
<tr>
<td></td>
<td>(0.0068)</td>
<td>(0.0107)</td>
<td>(0.0086)</td>
<td>(0.0138)</td>
</tr>
<tr>
<td>$(\delta_1 - \delta_2) - (\delta_1 - \delta_2)$</td>
<td>0.0959***</td>
<td>0.1215***</td>
<td>0.0716***</td>
<td>0.0499**</td>
</tr>
<tr>
<td></td>
<td>(0.0093)</td>
<td>(0.0146)</td>
<td>(0.0116)</td>
<td>(0.0187)</td>
</tr>
<tr>
<td><strong>Heterogeneity by gender of the parent for given level of parental education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_1 - \delta_1$</td>
<td>0.0576***</td>
<td>0.0633***</td>
<td>0.0481***</td>
<td>0.0152</td>
</tr>
<tr>
<td></td>
<td>(0.0102)</td>
<td>(0.0167)</td>
<td>(0.0125)</td>
<td>(0.0208)</td>
</tr>
<tr>
<td>$\delta_2 - \delta_2$</td>
<td>-0.0383***</td>
<td>-0.0582***</td>
<td>-0.0234**</td>
<td>-0.0348**</td>
</tr>
<tr>
<td></td>
<td>(0.0067)</td>
<td>(0.0119)</td>
<td>(0.0078)</td>
<td>(0.0142)</td>
</tr>
<tr>
<td>$\delta_3 - \delta_3$</td>
<td>0.0091*</td>
<td>0.0211**</td>
<td>0.0011</td>
<td>0.0200*</td>
</tr>
<tr>
<td></td>
<td>(0.0042)</td>
<td>(0.0080)</td>
<td>(0.0046)</td>
<td>(0.0092)</td>
</tr>
</tbody>
</table>

**Notes:** Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present differences between sets of OLS estimates of equation (10) in the paper, reported in Table 13. The outcome variable $STEM_{i2}$ denotes a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise; $FEdu_i, MEdu_{i,j}, i,j \in \{1, 2, 3, 4\}$ are dummy variables denoting the qualification level of fathers and mothers respectively, where $i = 1 (j = 1)$ if the father (mother) has STEM-degree qualification, $i = 2 (j = 2)$ if the father (mother) has a non-STEM degree qualification, $i = 3 (j = 3)$ if the father (mother) has a high school qualification and $i = 4 (j = 4)$ if the father (mother) has a junior high school (JHS) qualification or less. X includes the following covariates: region of residence, social class and parents’ jobs. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to differences between interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table Online Appendix - 8: Heterogeneity of the effects of parental education on STEM degree choice at the University, controlling for STEM high school, pooled and by gender of the student and of the parents. Interdependent effects.

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_{11} - \theta_{12}$</td>
<td>0.0474***</td>
<td>0.0574**</td>
<td>0.0490**</td>
<td>0.0084</td>
</tr>
<tr>
<td></td>
<td>(0.0140)</td>
<td>(0.0210)</td>
<td>(0.0185)</td>
<td>(0.0280)</td>
</tr>
<tr>
<td>$\theta_{11} - \theta_{13}$</td>
<td>0.0621***</td>
<td>0.0694***</td>
<td>0.0595***</td>
<td>0.0099</td>
</tr>
<tr>
<td></td>
<td>(0.0138)</td>
<td>(0.0208)</td>
<td>(0.0181)</td>
<td>(0.0276)</td>
</tr>
<tr>
<td>$\theta_{11} - \theta_{14}$</td>
<td>0.0951***</td>
<td>0.0935**</td>
<td>0.1079***</td>
<td>-0.0144</td>
</tr>
<tr>
<td></td>
<td>(0.0239)</td>
<td>(0.0373)</td>
<td>(0.0298)</td>
<td>(0.0477)</td>
</tr>
<tr>
<td>$\theta_{21} - \theta_{22}$</td>
<td>0.0245*</td>
<td>0.0308</td>
<td>0.0240</td>
<td>0.0068</td>
</tr>
<tr>
<td></td>
<td>(0.0109)</td>
<td>(0.0174)</td>
<td>(0.0135)</td>
<td>(0.0220)</td>
</tr>
<tr>
<td>$\theta_{21} - \theta_{23}$</td>
<td>0.0411***</td>
<td>0.0568***</td>
<td>0.0328*</td>
<td>0.0240</td>
</tr>
<tr>
<td></td>
<td>(0.0110)</td>
<td>(0.0177)</td>
<td>(0.0136)</td>
<td>(0.0223)</td>
</tr>
<tr>
<td>$\theta_{21} - \theta_{24}$</td>
<td>0.0492***</td>
<td>0.0844***</td>
<td>0.0301</td>
<td>0.0543</td>
</tr>
<tr>
<td></td>
<td>(0.0147)</td>
<td>(0.0245)</td>
<td>(0.0177)</td>
<td>(0.0302)</td>
</tr>
<tr>
<td>$\theta_{31} - \theta_{32}$</td>
<td>0.0821***</td>
<td>0.1011***</td>
<td>0.0745***</td>
<td>0.0265</td>
</tr>
<tr>
<td></td>
<td>(0.0122)</td>
<td>(0.0195)</td>
<td>(0.0150)</td>
<td>(0.0246)</td>
</tr>
<tr>
<td>$\theta_{31} - \theta_{33}$</td>
<td>0.0856***</td>
<td>0.0974***</td>
<td>0.0829***</td>
<td>0.0144</td>
</tr>
<tr>
<td></td>
<td>(0.0116)</td>
<td>(0.0186)</td>
<td>(0.0147)</td>
<td>(0.0235)</td>
</tr>
<tr>
<td>$\theta_{31} - \theta_{34}$</td>
<td>0.0847***</td>
<td>0.0785***</td>
<td>0.0877***</td>
<td>-0.0092</td>
</tr>
<tr>
<td></td>
<td>(0.0120)</td>
<td>(0.0195)</td>
<td>(0.0144)</td>
<td>(0.0245)</td>
</tr>
<tr>
<td>$\theta_{41} - \theta_{42}$</td>
<td>0.1246***</td>
<td>0.1425**</td>
<td>0.1244***</td>
<td>0.0181</td>
</tr>
<tr>
<td></td>
<td>(0.0280)</td>
<td>(0.0486)</td>
<td>(0.0330)</td>
<td>(0.0587)</td>
</tr>
<tr>
<td>$\theta_{41} - \theta_{43}$</td>
<td>0.1310***</td>
<td>0.1564*</td>
<td>0.1191***</td>
<td>0.0374</td>
</tr>
<tr>
<td></td>
<td>(0.0266)</td>
<td>(0.0460)</td>
<td>(0.0316)</td>
<td>(0.0558)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present differences between sets of OLS estimates on the pooled sample, on the sample of male students and on the sample of females students, respectively, in the same spirit of equation (3) in the paper, reported in Table 14. The outcome is $STEM_{i2}$ a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise. All equations include $X$, the following set of covariates: region of residence, social class and parents’ jobs. Estimates of column (4) correspond to differences between interaction terms of a fully interacted model specification by student gender. $\theta_{ij \in \{1, 2, 3, 4\}}$ denote the differential effect of having a father with qualification $i$ and a mother with qualification $j$ on the probability of completing a STEM University degree, respectively, with respect to the case in which both parents hold a qualification equal or lower to a junior high school degree. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table Online Appendix - 9: Estimates of the effects of parental education on STEM degree choice at the University, pooled and by gender of the student. Extended set of individual level controls, no high school mark.

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(mean: 0.231)</td>
<td>(mean: 0.348)</td>
<td>(mean: 0.155)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.1314***</td>
<td>0.1436***</td>
<td>0.1136***</td>
<td>0.0300**</td>
</tr>
<tr>
<td></td>
<td>(0.0057)</td>
<td>(0.0093)</td>
<td>(0.0071)</td>
<td>(0.0117)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>-0.0179***</td>
<td>-0.0413***</td>
<td>-0.0108**</td>
<td>-0.0305***</td>
</tr>
<tr>
<td></td>
<td>(0.0041)</td>
<td>(0.0072)</td>
<td>(0.0048)</td>
<td>(0.0086)</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.0143***</td>
<td>0.0168***</td>
<td>0.0078***</td>
<td>0.0090</td>
</tr>
<tr>
<td></td>
<td>(0.0025)</td>
<td>(0.0048)</td>
<td>(0.0028)</td>
<td>(0.0056)</td>
</tr>
<tr>
<td>Mother education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{1.1}$</td>
<td>0.0754***</td>
<td>0.0744***</td>
<td>0.0678***</td>
<td>0.0066</td>
</tr>
<tr>
<td></td>
<td>(0.0071)</td>
<td>(0.0114)</td>
<td>(0.0088)</td>
<td>(0.0144)</td>
</tr>
<tr>
<td>$\alpha_{1.2}$</td>
<td>0.0198***</td>
<td>0.0137*</td>
<td>0.0142***</td>
<td>-0.0005</td>
</tr>
<tr>
<td></td>
<td>(0.0041)</td>
<td>(0.0072)</td>
<td>(0.0048)</td>
<td>(0.0087)</td>
</tr>
<tr>
<td>$\alpha_{1.3}$</td>
<td>0.0055**</td>
<td>-0.0052</td>
<td>0.0063**</td>
<td>-0.0115**</td>
</tr>
<tr>
<td></td>
<td>(0.0026)</td>
<td>(0.0049)</td>
<td>(0.0029)</td>
<td>(0.0057)</td>
</tr>
<tr>
<td>$\alpha_{1. + \alpha_2.}$</td>
<td>0.1135***</td>
<td>0.1023***</td>
<td>0.1029***</td>
<td>-0.0005</td>
</tr>
<tr>
<td></td>
<td>(0.0083)</td>
<td>(0.0143)</td>
<td>(0.0100)</td>
<td>(0.0174)</td>
</tr>
<tr>
<td>$\alpha_{1. + \alpha_2}$</td>
<td>0.0952***</td>
<td>0.0881***</td>
<td>0.0820***</td>
<td>0.0061</td>
</tr>
<tr>
<td></td>
<td>(0.0094)</td>
<td>(0.0159)</td>
<td>(0.0113)</td>
<td>(0.0195)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of OLS estimates of equation (1) in the paper, reported here for convenience $STEM_{i2} = \alpha_0 + \alpha_1 F Edu_{1.} + \alpha_2 F Edu_{2.} + \alpha_3 F Edu_{3.} + \alpha_1 M Edu_{1.} + \alpha_2 M Edu_{2.} + \alpha_3 M Edu_{3.} \alpha_X X_i + \varepsilon_{i2}$, where $STEM_{i2}$ denotes a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise; $F Edu_{1.}, M Edu_{1.}, i, j \in \{1, 2, 3, 4\}$ are dummy variables denoting the qualification level of fathers and mothers respectively, where $i = 1 \ (j = 1)$ if the father (mother) has STEM-degree qualification, $i = 2 \ (j = 2)$ if the father (mother) has a non-STEM degree qualification, $i = 3 \ (j = 3)$ if the father (mother) has a high school qualification and $i = 4 \ (j = 4)$ if the father (mother) has a junior high school (JHS) qualification or less. $X$ includes the following covariates: region of residence, social class, parents’ jobs, delay in enrolment at University, delay in graduation from University, region of the University with respect to region of residence, academic and work values, type of high school. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table Online Appendix - 10: Estimates of the effects of parental education on STEM degree choice at the University, pooled and by gender of the student. Extended set of individual level controls, no high school mark. Interdependent effects.

<table>
<thead>
<tr>
<th>Pooled (1) (mean: 0.231)</th>
<th>Males (2) (mean: 0.348)</th>
<th>Females (3) (mean: 0.155)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_{11} )</td>
<td>0.2056*** (0.0120)</td>
<td>0.2113*** (0.0183)</td>
<td>0.1856*** (0.0158)</td>
</tr>
<tr>
<td>( \beta_{12} )</td>
<td>0.1570*** (0.0087)</td>
<td>0.1620*** (0.0138)</td>
<td>0.1330*** (0.0110)</td>
</tr>
<tr>
<td>( \beta_{13} )</td>
<td>0.1387*** (0.0080)</td>
<td>0.1443*** (0.0130)</td>
<td>0.1210*** (0.0099)</td>
</tr>
<tr>
<td>( \beta_{14} )</td>
<td>0.1047*** (0.0204)</td>
<td>0.1165*** (0.0326)</td>
<td>0.0798*** (0.0251)</td>
</tr>
<tr>
<td>( \beta_{21} )</td>
<td>0.0324*** (0.0108)</td>
<td>0.0119 (0.0176)</td>
<td>0.0289** (0.0133)</td>
</tr>
<tr>
<td>( \beta_{31} )</td>
<td>0.1084*** (0.0116)</td>
<td>0.1117*** (0.0188)</td>
<td>0.0925*** (0.0144)</td>
</tr>
<tr>
<td>( \beta_{41} )</td>
<td>0.1413*** (0.0261)</td>
<td>0.1694*** (0.0438)</td>
<td>0.1212*** (0.0313)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of OLS estimates of equation (3) in the paper on the pooled sample, on the sample of male students and on the sample of females students, respectively. The outcome is STEM\(_i\) a dummy taking the value 1 if student \( i \) graduates from a STEM field at University and 0 otherwise. All equations include \( X \), the following set of covariates: region of residence, social class, parents’ jobs, delay in enrolment at University, delay in graduation from University, region of the University with respect to region of residence, academic and work values, type of high school. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. \( \beta_{ij} i,j\in\{1,2,3,4\} \) denote the differential effect of having a father with qualification \( i \) and a mother with qualification \( j \) on the probability of completing a STEM University degree, respectively, with respect to the case in which both parents hold a qualification equal or lower to a junior high school degree. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table Online Appendix - 11: Estimates of the effects of parental education on STEM degree choice at the University, pooled and by gender of the student. Extended set of individual level controls, including high school mark.

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(mean: 0.213)</td>
<td>(mean: 0.348)</td>
<td>(mean: 0.155)</td>
<td></td>
</tr>
<tr>
<td>Father education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.1232***</td>
<td>0.1330***</td>
<td>0.1063***</td>
<td>0.0267**</td>
</tr>
<tr>
<td></td>
<td>(0.0057)</td>
<td>(0.0091)</td>
<td>(0.0070)</td>
<td>(0.0115)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>-0.0220***</td>
<td>-0.0448***</td>
<td>-0.0158***</td>
<td>-0.0290***</td>
</tr>
<tr>
<td></td>
<td>(0.0041)</td>
<td>(0.0070)</td>
<td>(0.0048)</td>
<td>(0.0085)</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.0133***</td>
<td>0.0151***</td>
<td>0.0067**</td>
<td>0.0083</td>
</tr>
<tr>
<td></td>
<td>(0.0025)</td>
<td>(0.0047)</td>
<td>(0.0028)</td>
<td>(0.0055)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.0648***</td>
<td>0.0609***</td>
<td>0.0590***</td>
<td>0.0019</td>
</tr>
<tr>
<td></td>
<td>(0.0070)</td>
<td>(0.0111)</td>
<td>(0.0087)</td>
<td>(0.0141)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.0135***</td>
<td>0.0075</td>
<td>0.0079*</td>
<td>-0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0041)</td>
<td>(0.0071)</td>
<td>(0.0048)</td>
<td>(0.0085)</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.0037</td>
<td>-0.0043</td>
<td>0.0036</td>
<td>-0.0080</td>
</tr>
<tr>
<td></td>
<td>(0.0026)</td>
<td>(0.0048)</td>
<td>(0.0028)</td>
<td>(0.0056)</td>
</tr>
<tr>
<td>$\alpha_1 + \alpha_2$</td>
<td>0.1012***</td>
<td>0.0882***</td>
<td>0.0905***</td>
<td>-0.0023</td>
</tr>
<tr>
<td></td>
<td>(0.0082)</td>
<td>(0.0139)</td>
<td>(0.0099)</td>
<td>(0.0171)</td>
</tr>
<tr>
<td>$\alpha_1 + \alpha_2$</td>
<td>0.0783***</td>
<td>0.0684***</td>
<td>0.0669***</td>
<td>0.0015</td>
</tr>
<tr>
<td></td>
<td>(0.0093)</td>
<td>(0.0155)</td>
<td>(0.0112)</td>
<td>(0.0192)</td>
</tr>
</tbody>
</table>

Notes: Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of OLS estimates of equation (1) in the paper. The outcome is $STEM_{i2}$ denotes a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise; $F_{Edu,i}$, $M_{Edu,j}$, $i,j \in \{1,2,3,4\}$ are dummy variables denoting the qualification level of fathers and mothers respectively, where $i = 1$ ($j = 1$) if the father (mother) has STEM-degree qualification, $i = 2$ ($j = 2$) if the father (mother) has a non-STEM degree qualification, $i = 3$ ($j = 3$) if the father (mother) has a high school qualification and $i = 4$ ($j = 4$) if the father (mother) has a junior high school (JHS) qualification or less. $X$ includes the following covariates: region of residence, social class, parents’ jobs, delay in enrolment at University, delay in graduation from University, region of the University with respect to region of residence, academic and work values, type of high school and high school graduation mark. Column (1) uses data for all students. Column (2) and (3) use data on male and female students only, respectively. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.
Table Online Appendix - 12: Estimates of the effects of parental education on STEM degree choice at the University, pooled and by gender of the student. Extended set of individual level controls, including high school mark. Interdependent effects.

<table>
<thead>
<tr>
<th></th>
<th>Pooled (1)</th>
<th>Males (2)</th>
<th>Females (3)</th>
<th>Males-Females (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(mean: 0.231)</td>
<td>(mean: 0.348)</td>
<td>(mean: 0.155)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Both parents with degree in STEM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>0.1856***</td>
<td>0.1857***</td>
<td>0.1679***</td>
<td></td>
</tr>
<tr>
<td>(0.0119)</td>
<td>(0.0178)</td>
<td>(0.0157)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Father with degree in STEM different mother education levels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{12}$</td>
<td>0.1424***</td>
<td>0.1447***</td>
<td>0.1197***</td>
<td></td>
</tr>
<tr>
<td>(0.0086)</td>
<td>(0.0136)</td>
<td>(0.0109)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{13}$</td>
<td>0.1295***</td>
<td>0.1360***</td>
<td>0.1116***</td>
<td></td>
</tr>
<tr>
<td>(0.0079)</td>
<td>(0.0127)</td>
<td>(0.0099)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{14}$</td>
<td>0.1033***</td>
<td>0.1183***</td>
<td>0.0768***</td>
<td></td>
</tr>
<tr>
<td>(0.0203)</td>
<td>(0.0328)</td>
<td>(0.0248)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mother with degree in STEM different father education levels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>0.0187*</td>
<td>-0.0045</td>
<td>0.0159</td>
<td></td>
</tr>
<tr>
<td>(0.0107)</td>
<td>(0.0173)</td>
<td>(0.0132)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{31}$</td>
<td>0.0990***</td>
<td>0.1011***</td>
<td>0.0839***</td>
<td></td>
</tr>
<tr>
<td>(0.0115)</td>
<td>(0.0184)</td>
<td>(0.0142)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{41}$</td>
<td>0.1279***</td>
<td>0.1472***</td>
<td>0.1120***</td>
<td></td>
</tr>
<tr>
<td>(0.0257)</td>
<td>(0.0423)</td>
<td>(0.0309)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** Source: Almalaurea XIX Profilo dei Laureati survey (2016 cohort of graduates). Sample: 155,603 students who graduated from high school and from University in Italy, from 3-years and 5-years degree cycles with non-missing value of all covariates included in the analysis. Columns (1)-(3) present different sets of OLS estimates of equation (3) in the paper on the pooled sample, on the sample of male students and on the sample of females students, respectively. The outcome is $STEM_{ij}$ a dummy taking the value 1 if student $i$ graduates from a STEM field at University and 0 otherwise. All equations include $X$, the following set of covariates: region of residence, social class, parents’ jobs, delay in enrolment at University, delay in graduation from University, region of the University with respect to region of residence, academic and work values, type of high school and high school graduation mark. Estimates of column (4) correspond to interaction terms of a fully interacted model specification by student gender. $\beta_{ij}$, $i,j \in \{1, 2, 3, 4\}$ denote the differential effect of having a father with qualification $i$ and a mother with qualification $j$ on the probability of completing a STEM University degree, respectively, with respect to the case in which both parents hold a qualification equal or lower to a junior high school degree. Robust standard errors in parentheses. * significant at 5%; ** significant at 1% or better.